Erzincan Üniversitesi
Fen Bilimleri Enstitüsü Dergisi
2024, 17(3), 779-788
ISSN: 1307-9085, e-ISSN: 2149-4584
Araştırma Makalesi

Erzincan University
Journal of Science and Technology
2024, 17(3), 779-788
DOI: 10.18185/erzifbed.1552843
Research Article

Block-Based Forgery Detection: Performance Comparison Using HOG, LBP and MBF

Yıldız AYDIN[1*], Yunus BABACAN[2]

[1]Department of Computer Engineering, Erzincan Binali Yıldırım University, Erzincan, 24000
[2]Department of Electrical and Electronics Engineering, Erzincan Binali Yıldırım University, Erzincan, 24000

Received: 19/09/2024, **Revised:** 23/10/2024, **Accepted:** 07/11/2024, **Published:** 31/12/2024

Abstract

One of the types of forgery performed on digital images is copy-move forgery (CMF). This type of forgery is carried out by pasting a region copied from the same image over another region of the image. It is very important to determine whether there is any forgery on these images, as they can be used as evidence in many fields. In this study, an analysis of forgery detection is performed using Histogram of Oriented Gradients (HOG), Local Binary Patterns (LBP), and Multiscale Basic Features (MBF) for block-based copy-move forgery detection. The performance of various features, both individually and in combination, is evaluated. Combinations such as HOG+LBP, HOG+MBF, and MBF+LBP were tested, but the expected performance improvement was not achieved. Although the performance increase was not significant, the highest results were generally obtained with the LBP+MBF hybrid feature, resulting in an F1 score of 88.5%. Additionally, while HOG and LBP features are frequently used in block-based approaches, the use of the MBF feature has not been found in the literature. This study contributes to the existing methods in the field of block-based forgery detection and highlights the effectiveness of various features and feature combinations.

Keywords: Copy-move forgery, HOG, LBP, MBF

Blok Tabanlı Sahtecilik Tespiti: HOG, LBP ve MBF Kullanılarak Performans Karşılaştırması

Öz

Dijital görüntüler üzerinde yapılan sahtecilik türlerinden biri de kopyala-yapıştır sahteciliğidir (CMF). Bu sahtecilik türü, aynı görüntüden kopyalanan bir bölgenin, görüntünün başka bir bölgesi üzerine yapıştırılmasıyla gerçekleştirilir. Bu görüntüler, birçok alanda delil olarak kullanılabileceğinden, üzerinde herhangi bir sahtecilik olup olmadığının belirlenmesi oldukça önemlidir. Bu çalışmada, blok tabanlı kopyala-yapıştır sahteciliği tespiti için Yönlendirilmiş Gradyanların Histogramı (HOG), Yerel İkili Örüntüler (LBP) ve Çok Ölçekli Temel Öznitelikler (MBF) kullanılarak sahtecilik tespiti üzerine bir analiz yapılmıştır. Çeşitli özniteliklerin hem tek başına hem de birlikte performansları değerlendirilmiştir. HOG+LBP, HOG+MBF ve MBF+LBP gibi kombinasyonlar denenmiş, ancak beklenen performans artışı sağlanamamıştır. Performans artışı çok büyük olmasa da, en yüksek sonuçlar genellikle LBP+MBF hibrit özelliği ile elde edilmiş ve %88,5'lik bir F1 puanı ile sonuçlanmıştır. Ayrıca, HOG ve LBP özellikleri blok tabanlı yaklaşımlarda sıkça kullanılsa da, MBF özniteliğini kullanan yaklaşımlara literatürde rastlanmamıştır. Bu çalışma, blok tabanlı sahtecilik tespiti alanındaki mevcut yöntemlere katkı sağlamakta ve çeşitli öznitelik ve öznitelik kombinasyonlarının etkinliğini vurgulamaktadır.

Anahtar Kelimeler: Kopyala yapıştır sahteciliği, HOG, LBP, MBF.

*Corresponding Author: yciltas@erzincan.edu.tr
Yıldız AYDIN, https://orcid.org/0000-0002-3877-6782
Yunus BABACAN, https://orcid.org/0000-0002-6745-0626

1. Introduction

The rapid development of technology has significantly increased the use of digital images and the operations performed on these images. Along with this increase, studies on forgery detection systems on digital images have become increasingly important. In particular, the widespread use and ease of use of image editing applications such as Photoshop has led to a rapid increase in the number of forged images, making it crucial to accurately and reliably detect whether such images are forged or not, as they can be used as evidence in many areas, such as forensic cases.

To make it difficult to detect forgeries, various post-processing operations such as blurring, noise addition, and illumination modification are performed on forged images. This makes forgery detection more challenging and increases the need for the development of a reliable copy-move forgery detection system. Therefore, the focus of much research in this area has been the development of robust and effective detection methods against forged images.

Methods used in copy-move forgery detection (CMFD) systems are generally divided into two groups: keypoint-based methods and block-based methods. Keypoint-based methods have the disadvantage that the boundaries of the forged region cannot be precisely determined. In block-based methods, the entire image information is used, but this can negatively affect the accuracy rate. In the literature, various studies have been conducted in order to improve the performance of copy-move forgery (CMF) detection systems, but research on the use of hybrid features in block-based approaches has been very limited.

In the field of image processing, the use of hybrid features has been widely investigated in keypoint-based methods, where successful results have been achieved [1–3]. However, the lack of emphasis on the potential benefits of hybrid features in block-based approaches points to an important gap in this field. The main objective of this study is to investigate the use of hybrid features in block-based CMFD systems to address the challenges of detecting forgery region boundaries and to present a new method aimed at improving performance.

Main Contributions of the Proposed Method to the Literature:

- The proposed method introduces a new CMFD (Copy-Move Forgery Detection) approach that compares the use of HOG, LBP, MBF, and hybrid features in block-based forgery detection systems.

- The problem of not being able to detect forgery zone boundaries, which is frequently encountered in keypoint-based methods, has been solved.

In this research, we propose a CMFD system that utilizes various hybrid features in a block-based approach. Thus, the aim is to compare the performance of these features while addressing the disadvantage of not being able to detect forgery boundaries, which is often encountered in keypoint-based methods. The rest of this paper is organized as follows: In the second section of this article, the related work on copy-move forgery detection methods are detailed. In the third section, the stages of the proposed method are explained. In the fourth section, experimental results and comparisons of the proposed method with other methods in the literature are presented. Finally, in the fifth section, the results of the study and future studies are detailed.

2. Related Work

Copy-move forgery detection is a popular area of research that is being intensively studied. The methods developed for forgery detection in digital images with evidentiary value can be basically divided into 3 categories. These categories are keypoint-based, block-based and hybrid approaches.

In key-point based CMFD systems, in the first stage, identifiers are created on the key points detected on the image. The keypoint-based descriptors (local features) obtained from the whole image are compared with each other. As a result of this comparison, if there are local features that match each other, the locations of these features are marked as forged regions. Amerini et al. [4] used the SIFT feature, which is invariant to post-processing such as scaling differences, illumination differences and rotation changes, in their proposed CMFD system. In the application, forgery detection was performed by taking into account the matches between the extracted SIFT features.

In recent years, many studies have been conducted to increase the success rate by using hybrid features [1,5,6]. In the study by Aydın [7] keypoint locations are detected using the DOG detector and a LIOP descriptor is built on these keypoints. Thanks to this hybrid feature, higher performance was achieved on different datasets. In the method proposed by Wang et al. [8], simple linear iterative clustering (SLIC) and the K-multiple-means methods are used to detect keypoint locations. They obtained a hybrid feature using Fast Quaternion Generic Polar Complex Exponential Transform (FQGPCET) and Graylevel co-occurrence matrix (GLCM) as features. They obtained a high precision value in the experimental results of their proposed method.

In block-based approaches, the image is decomposed into fixed-size blocks and features are extracted from each block, and forgery defects persist throughout these features. Fridrich et al. [9] were among the pioneers in this area, proposing a method where the image is divided into blocks, and Discrete Cosine Transform (DCT) feature vectors are extracted. These vectors are then lexically filtered and compared to detect forgeries. Ganguly et al. [10] suggested an approach where local tetra pattern-based texture descriptors are extracted from the blocks, which are then compared to identify forgery. This method has shown increased robustness to common post-processing and has been particularly effective in detecting small forgeries that traditional methods miss. Shehin et al. [11] present a method that combines the Discrete Cosine Transform (DCT) with eigenvalues to enhance the detection and localization of forgery, especially in images subjected to post-processing operations like rotation. This method processes overlapping blocks of the image, extracting features using DCT and eigenvalues, and applies cumulative DCT features for more robust detection, even against rotation attacks. Weng et al. [12] proposed a new CMFD method based on UCM-Net, a U-Net-like architecture for the problem of forgery detection, which is difficult to solve due to the photometric similarities between forged regions and original regions. In particular, the method treats large and small forged regions differently and uses various deep networks and techniques to extract the features of these regions and localize the forged regions more precisely. UCM-Net provided higher accuracy rates in the face of operations such as blurring, rotation, and noise that make forgery detection difficult. Experiments have shown that UCM-Net outperforms existing best practices.

3. The Proposed Method

This paper proposes a three-stage method for the detection of copy-move forgery: (i) decomposition of the image into overlapping blocks, (ii) feature extraction and fusion of the extracted features, (iii) detection of forgery regions. There are several studies in the literature that apply various histogram-based local features using block-based approaches [10,13]. However, the combination of different feature extraction techniques such as Local Binary Pattern (LBP), Histogram of Oriented Gradients (HOG) and Multiscale Basic Features (MBF) and performance comparisons are limited. In Aydın's skin cancer recognition study [14], the combination of these features has been shown to yield better results.

In recent years, keypoint-based features have been widely used in copy-move forgery [15,16], and hybrid methods developed by combining different keypoint based methods are becoming widespread [1, 8]. In this study, the performance comparison of the features used in block-based approaches and the fusion features obtained by combining these features in various combinations was performed. Figure 1 presents the flowchart of the proposed method, which shows the general operation of the proposed method.

In the first stage of the proposed method, the analyzed image is divided into overlapping 8x8 blocks. The 8x8 block size is chosen because it is widely used in block-based approaches in the literature [10, 18]. While smaller blocks require higher computational costs and increased memory consumption, larger blocks may result in a loss of detail and decreased detection accuracy.

In the second stage, Local Binary Pattern (LBP), Histogram of Oriented Gradients (HOG) and Multiscale Basic Features (MBF) are extracted separately from each block. For each of these methods, 9-dimensional feature vectors are obtained from LBP and HOG, and 1536-dimensional feature vectors are obtained from MBF. In order to evaluate the performance of hybrid features, these individual features are combined to form new feature sets. For example, the HOG+LBP combination provides 18-dimensional feature vectors, HOG+MBF provides 1545-dimensional feature vectors and LBP+MBF provides 1545-dimensional feature vectors. These hybrid features are named as HOG+LBP, HOG+MBF and LBP+MBF, respectively.

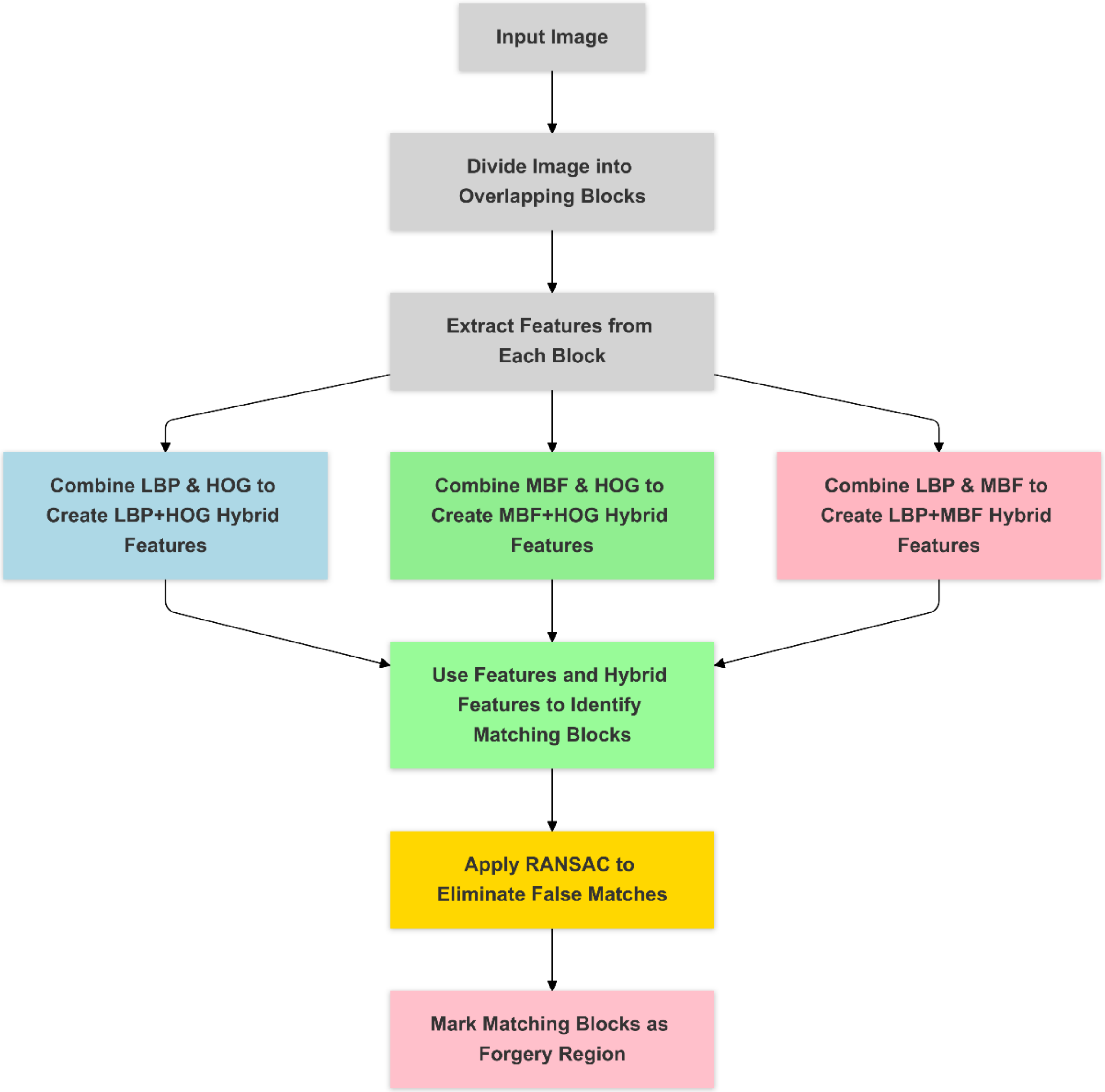
Input Image
Divide Image into Overlapping Blocks
Extract Features from Each Block
Combine LBP & HOG to Create LBP+HOG Hybrid Features
Combine MBF & HOG to Create MBF+HOG Hybrid Features
Combine LBP & MBF to Create LBP+MBF Hybrid Features
Use Features and Hybrid Features to Identify Matching Blocks
Apply RANSAC to Eliminate False Matches
Mark Matching Blocks as Forgery Region

Figure 1. Flowchart of the suggested approach

In the last stage, all extracted features in the second stage are stored in the M matrix and sorted according to the dictionary order. If the Euclidean distance between each feature vector and the other vectors in the M matrix is less than the previously determined threshold value, these vectors are considered to be matched and the relevant blocks are labeled as copy-move forgery regions. In this method, hybrid features obtained by combining the features frequently used in keypoint-based methods are used in the block-based approach.

3.1 Dataset

The CoMoFoD dataset is a rich database that is widely used in research on digital image forgery and contains different forms of forgery. For each 40 images in the CoMoFoD dataset, a total of 200 images were created by applying operations such as translation, rotation, scaling, distortion, combination. In addition, different types of attacks such as image blurring (IB), color reduction (CR), JPEG-compression, contrast adjustment (CA) and changes in brightness (BC) were applied to these 200 images in the CoMoFoD dataset. The final number of all these 512x512 images is 10,400. In this study, only the first 40 images containing translation in the CoMoFoD (Copy-Move Forgery Detection) dataset were used to measure the performance of the application developed for copy-move forgery detection.

4. Experimental Results

In this study, experiments were conducted to compare the performance of various features for the copy-move forgery detection problem. The experiments were performed on a computer with an Intel i5 processor and 8 GB of RAM, using the Python programming language. In this context, HOG, LBP, MBF and combinations such as HOG+LBP, HOG+MBF, and LBP+MBF were included as features.

The experiments were conducted on 40 plain copy-move forgery applied images with a resolution of 512x512 from the CoMoFoD dataset. Precision, recall, F1-score, and accuracy were used as performance metrics. Precision measures the ratio of pixels that are actually fake among all pixels labeled as fake, while recall measures the ratio of correctly predicted fake pixels to the total fake pixels. In short, precision focuses on the false positives, while recall focuses on the rate of forgery detection. Since both metrics are important, the harmonic mean of precision and recall is calculated as the F1-score. Lastly, the accuracy is calculated as the ratio of correctly labeled pixels to the total number of pixels. The calculation formulas for precision, recall, F1 score and accuracy metrics are given in Equations 1, 2, 3 and 4, respectively.

$$recall(r_c) = \frac{TP}{TP+FN} \tag{1}$$

$$Precision(p_r) = \frac{TP}{TP+FP} \tag{2}$$

$$F1\,Score = 2 * \frac{Recall * Precision}{Recall + Precision} \tag{3}$$

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \tag{4}$$

Table 1 shows the results of the experiments performed using 40 plain copy-move forgery images from the CoMoFoD dataset.

Table 1. Results of experiments

Method	Precision	Recall	F1 Score	Accuracy
HOG	0.297	**0.928**	0.451	0.799
LBP	0.868	0.894	0.881	0.986
MBF	**0.895**	0.869	0.882	**0.990**

Method	**Precision**	**Recall**	**F1 Score**	**Accuracy**
HOG+LBP	0.879	0.871	0.875	0.989
HOG+MBF	0.874	0.873	0.874	0.989
LBP+MBF	0.878	0.891	**0.885**	**0.990**

As shown in Table 1, the experiment using the HOG feature resulted in a high recall value but a low precision value. This shows that the HOG feature tends to over-detect fake regions, which leads to a large number of false positives.

Table 2 presents the performance comparison of the proposed method with other methods in the literature. The performance results in Table 2 are obtained using 40 plain copy-move forged images from the CoMoFoD dataset. Except for the proposed method, other method results are taken from [10]

Table 2. Comparative analysis for plain copy-move forgery applied to 40 forged images in CoMoFoD dataset.

Method	**Year**	**Precision**	**Recall**	**F1 Score**
Li et al. [18]	2014	0.418	0.833	0.557
Silva et al. [19]	2015	0.492	0.775	0.602
Liu et al. [20]	2018	0.455	0.802	0.581
Kumar et al. [21]	2022	**0.910**	0.818	0.844
Proposed Method		0.878	**0.891**	**0.885**

The precision and recall values obtained in the experiments performed using the LBP and MBF features are consistent and higher. These methods have shown an effective performance in detecting fake and non-fake regions. In the experiments conducted with the HOG+LBP, HOG+MBF, LBP+MBF methods obtained by the combination of features, they showed slightly better performance than the individual methods. The best overall performance was obtained by LBP+MBF with 0.878 precision, 0.891 recall and 0.990 accuracy.

In Figure 2, some prediction images obtained from the experiment performed with the LBP+MBF features, which obtained the highest performance, are given.

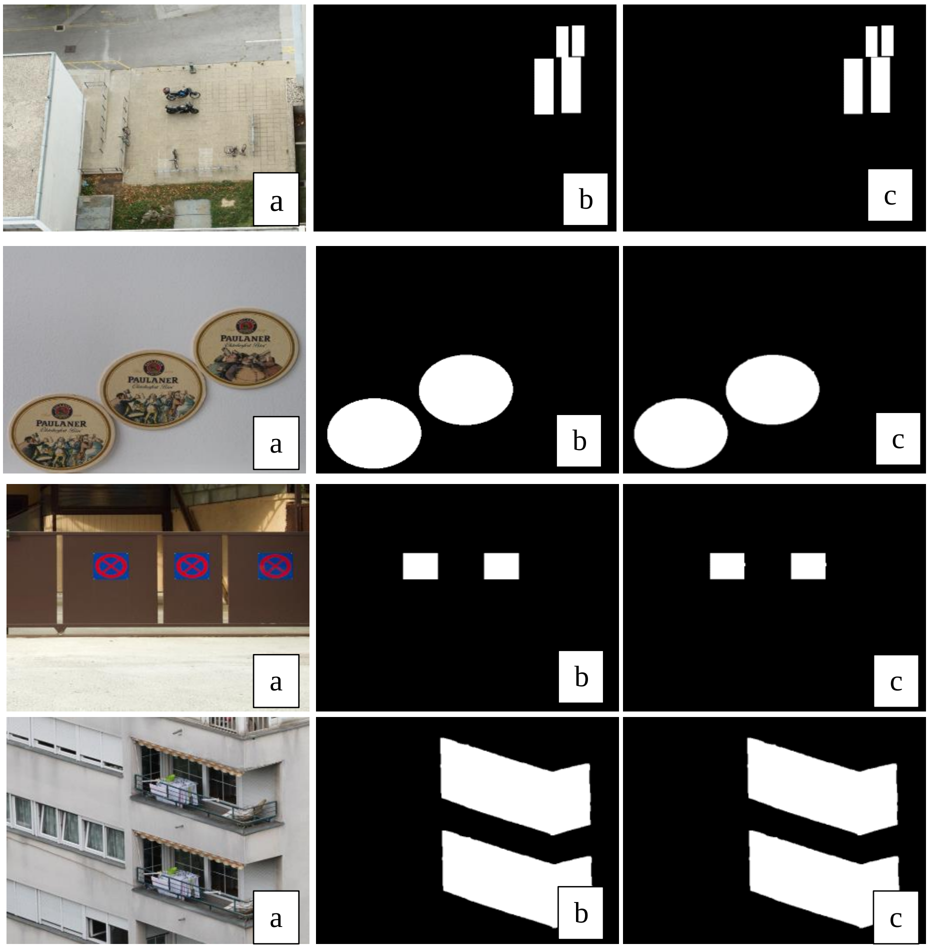

Figure 2. Results of the suggested approach on plain copy-move forgery images from the CoMoFoD dataset: a) counterfeit image, b) ground truth, c) detection result of the proposed method.

5. Conclusions

In this study, the performances obtained by using HOG, LBP, MBF features and different combinations of these features in the block-based approach were compared for the copy-move forgery detection problem. The experiments were carried out by measuring the similarities of HOG, LBP, MBF and HOG+LBP, HOG+MBF and LBP+MBF features on 40 forged images in the CoMoFoD dataset. In the performance evaluation performed using precision, recall, f1-score and accuracy evaluation metrics, the application performed with LBP+MBF provided the highest performance. The lowest performance was obtained in the experiment performed using the HOG feature. In the future, the use of more robust features and machine learning models will be investigated to improve the detection rate in the block-based approach.

Ethics in Publishing

There are no ethical issues regarding the publication of this study.

Author Contributions

All authors contributed equally to all stages of the study, including its conceptualization, methodology, data collection, analysis, and writing.

Acknowledgements

This project was supported by the Scientific Research Projects (BAP) program under project number 1004 FBA-2024-1004, titled 'Block-based Copy-move forgery Detection,' and conducted at Erzincan Binali Yıldırım University"

References

[1] Niyishaka, P., Bhagvati, C. (2020). Copy-move forgery detection using image blobs and BRISK feature. Multimed. Tools Appl. 10.1007/s11042-020-09225-6.

[2] Aydın, Y. (2024). Automated identification of copy-move forgery using Hessian and patch feature. J. Forensic Sci. *69*, 131–138.

[3] Sunitha, K., Krishna, A.N., Prasad, B.G. (2022). Copy-move tampering detection using keypoint based hybrid feature extraction and improved transformation model. Appl. Intell., 15405–15416. 10.1007/s10489-022-03207-x.

[4] Amerini, I., Ballan, L., Caldelli, R., Del Bimbo, A., Serra, G. (2011). A SIFT-based forensic method for copy-move attack detection and transformation recovery. IEEE Trans. Inf. Forensics Secur. *6*, 1099–1110. 10.1109/TIFS.2011.2129512.

[5] Kumar, N., Meenpal, T. (2022). Salient keypoint-based copy-move image forgery detection. 10.1080/00450618.2021.2016964.

[6] Aydin, Y. (2022). Comparison of color features on copy-move forgery detection problem using HSV color space. Aust. J. Forensic Sci. *early acce*. 10.1080/00450618.2022.2157046.

[7] Aydın, Y. (2022). A new Copy-Move forgery detection method using LIOP. J. Vis. Commun. Image Represent. *89*, 103661. 10.1016/j.jvcir.2022.103661.

[8] Wang, X. yang, Wang, X. qi, Niu, P. pan, Yang, H. ying (2024). Accurate and robust image copy-move forgery detection using adaptive keypoints and FQGPCET-GLCM feature (Springer US) 10.1007/s11042-023-15499-3.

[9] Fridrich, J., Soukal, D., Lukáš, J. (2003). Detection of Copy-Move Forgery in Digital Images. Proc. Digit. Forensic Res. Work., 133–162. 10.1109/ICMLA.2015.137.

[10] Ganguly, S., Mandal, S., Malakar, S., Sarkar, R. (2023). Copy-move forgery detection using local tetra pattern based texture descriptor. Multimed. Tools Appl. *82*, 19621–19642. 10.1007/s11042-022-14287-9.

[11] Shehin, A.U., Sankar, D. (2024). Copy Move Forgery detection and localisation robust to rotation using block based Discrete Cosine Transform and eigenvalues. J. Vis. Commun. Image Represent. *99*, 104075. 10.1016/j.jvcir.2024.104075.

[12] Weng, S., Zhu, T., Zhang, T., Zhang, C. (2024). UCM-Net: A U-Net-Like Tampered-Region-Related Framework for Copy-Move Forgery Detection. IEEE Trans. Multimed. *26*, 750–763. 10.1109/TMM.2023.3270629.

[13] Nawaz, M., Mehmood, Z., Nazir, T., Masood, M., Tariq, U., Munshi, A.M., Mehmood, A., Rashid, M. (2021). Image authenticity detection using DWT and circular block-based LTrP features. Comput. Mater. Contin. *69*, 1927–1944. 10.3233/JIFS-191700.

[14] Aydin, Y. (2023). A Comparative Analysis of Skin Cancer Detection Applications Using Histogram-Based Local Descriptors. Diagnostics *13*. 10.3390/diagnostics13193142.

[15] Hailing, H., Weiqiang, G., Yu, Z. (2008). Detection of copy-move forgery in digital images using sift algorithm. Proc. - 2008 Pacific-Asia Work. Comput. Intell. Ind. Appl. PACIIA 2008 *2*, 272–276. 10.1109/PACIIA.2008.240.

[16] Raj, R., Joseph, N. (2016). Keypoint Extraction Using SURF Algorithm for CMFD. Procedia Comput. Sci. *93*, 375–381. 10.1016/j.procs.2016.07.223.

[17] Popescu, A.C., Farid, H. (2004). Exposing Digital Forgeries by Detecting Duplicated Image Regions. Tech. Report, TR2004-515, Dep. Comput. Sci. Dartmouth Coll. Hanover, New Hampsh., 1–11.

[18] Li, J., Li, X., Yang, B., Sun, X. (2015). Segmentation-based image copy-move forgery detection scheme. IEEE Trans. Inf. Forensics Secur. *10*, 507–518. 10.1109/TIFS.2014.2381872.

[19] Silva, E., Carvalho, T., Ferreira, A., Rocha, A. (2015). Going deeper into copy-move forgery detection: Exploring image telltales via multi-scale analysis and voting processes. J. Vis. Commun. Image Represent. *29*, 16–32. 10.1016/j.jvcir.2015.01.016.

[20] Liu, Y., Guan, Q., Zhao, X. (2018). Copy-move forgery detection based on convolutional kernel network. Multimed. Tools Appl. *77*, 18269–18293. 10.1007/s11042-017-5374-6.

[21] Kumar, S., Mukherjee, S., Pal, A.K. (2023). An improved reduced feature-based copy-move forgery detection technique. Multimed. Tools Appl. *82*, 1431–1456. 10.1007/s11042-022-12391-4.