Journal of Artificial Intelligence and Data Mining (JAIDM), Vol. 11, No. 3, 2023, 429-442.

Shahrood University of Technology

Journal of Artificial Intelligence and Data Mining (JAIDM)
Journal homepage: http://jad.shahroodut.ac.ir

Research paper

An Optimal Hybrid Method to Detect Copy-move Forgery

Fatemeh Zare Mehrjardi[1], Ali Mohammad Latif[1*] and Mohsen Sardari Zarchi[2]

1. Computer Engineering Department, Yazd University, Yazd, Iran.
2. Computer Engineering Department, Meybod University, Meybod, Yazd, Iran.

Article Info

Article History:
Received 25 May 2023
Revised 20 July 2023
Accepted 19 August 2023

DOI:10.22044/jadm.2023.13166.2453

Keywords:
Copy-move Forgery, Block-based Method, Keypoint-based Method, Hybrid Method, Genetic Algorithm, Simulating Annealing Algorithm.

**Corresponding author: alatif@yazd.ac.ir (A. M. Latif).*

Abstract

Image is a powerful communication tool that is widely used in various applications such as forensic medicine and court, where the validity of the image is crucial. However, with the development and availability of image editing tools, image manipulation can be easily performed for a specific purpose. Copy-move forgery is one of the simplest and most common methods of image manipulation. There are two traditional methods to detect this type of forgery: block-based and key point-based. In this study, we present a hybrid approach of block-based and key point-based methods using meta-heuristic algorithms to find the optimal configuration. For this purpose, we first search for pair blocks suspected of forgery using the genetic algorithm with the maximum number of matched key points as the fitness function. Then we find the accurate forgery blocks using simulating annealing algorithm and producing neighboring solutions around suspicious blocks. We evaluate the proposed method on CoMoFod and COVERAGE datasets, and obtain the results of accuracy, precision, recall, and IoU with values of 96.87, 92.15, 95.34, and 93.45, respectively. The evaluation results show the satisfactory performance of the proposed method.

1. Introduction

Images are among the most powerful communication tools between humans. Digital devices such as cameras and cell phones make image generation easy at any time and place. In some applications, images can serve as evidence. However, if these images are manipulated, they will lose their credibility. Manipulation is done to hide or add information to the image, creating a forgery. In forgery images, the structure and texture of the images are altered [1].

Forgery images can be generated by two major approaches: active and passive. In the active approach, information is inserted into the original image to create a forgery image. This approach requires both the original and the forgery images to extract the information from the forgery image. Examples of active approaches are digital watermarking and digital signatures [2].

In the passive approach, forgery images are created by inserting, removing or modifying parts of images. This approach does not need any prior information such as the original image, unlike the active approach, so it is more popular. Passive approaches include copy-move, image splicing, image retouching, and object removal [3].

Copy-move forgery is one of the most common methods of creating a forgery image. It involves copying and pasting one or more regions of an image into other regions of the same image. To make the forgery more realistic, geometric transformations and post-processing operations are also applied among with copying and pasting. This method is easy to implement but hard to detect [4].

Traditional copy-move forgery detection methods can be mainly classified into two categories: block-based and key-point-based methods [5].

In the block-based method, first in the pre-processing step, an image is divided into overlapping or non-overlapping rectangular or circular blocks. Next, in the feature extraction step,

feature vectors are extracted from all blocks by feature extraction algorithms such as Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT), Local Binary Patterns (LBP), Polar Complex Exponential Transform (PCET), etc. [5-9]. Finally, in the feature matching step, similar blocks are found using sorting, correlation, and calculating Euclidean distance between feature vectors. The block-based method is easy to implement, but it has high computational complexity and poor performance against geometric transformations such as rotation and scaling. [10].In the key-point-based method, in the feature extraction step, feature points are extracted from the whole image using various key point extraction algorithms such as Scale-Invariant Feature Transform (SIFT) [11], Speeded Up Robust Features (SURF) [12], Binary Robust Invariant Scalable Keypoints (BRISK) [13], and Features from Accelerated Segment Test (FAST) [14] algorithms without dividing the image. In the feature-matching step, key points are matched based on their feature vectors with different approaches such as clustering, Euclidean distance, and nearest neighborhood.

The key-point-based method has low computational complexity and suitable performance against geometric transformations and post-processing operations. However, poor performance in detecting small and smooth forgery regions due to the lack of sufficient key points is one disadvantage of this method. A summary of the mentioned contents is given in Figure 1.

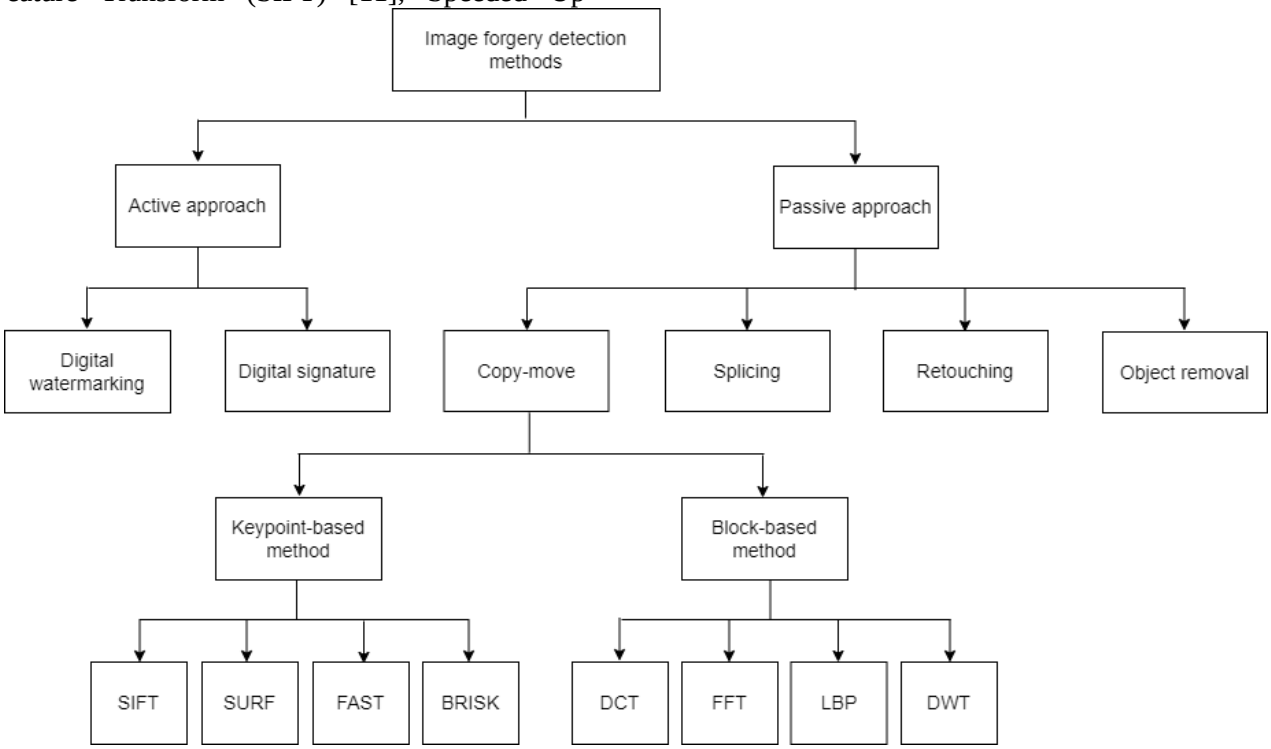

Figure 1. Summary of copy-move forgery detection techniques.

This study focuses on copy-move forgery detection and presents a hybrid method that combines block-based and key-point-based methods, and meta-heuristic algorithms. The r

est of this study is organized as what follows. Section 2 provides a brief overview of previous research works in forgery detection. Section 3 explains three meta-heuristic algorithms that were used in this study. Section 4 describes the details of the proposed method. Section 5 evaluates the experiment results and compares them with other methods. Section 6 concludes the study.

2. Related works

In the recent years, different forgery detection algorithms have been performed based on three methods, block-based, key-point-based, and hybrid methods. This section reviews some studies on these methods. For example, Mahmood *et al*. in 2016 proposed a block-based method that used a combination of DCT and Principal Component Analysis (PCA) algorithms to detect copy-move forgery [15]. They converted the RGB image into a gray image in the pre-processing step. Then they split the gray image into overlapping square blocks. In the feature extraction step, they used the DCT components and extracted feature vectors from all blocks. These feature vectors had high dimensional feature space, so they applied the PCA algorithm to achieve the reduced dimensional feature vector

representation. Finally, they found similar blocks by calculating the Euclidean distance between all block features.

The color and texture information are two important components for copy-move forgery detection. Zhu *et al*. proposed a block-based method that used these two components with the color local binary patterns (CoLBP) algorithm [16]. They applied the CoLBP algorithm to the image in the pre-processing step to combine the color information and LBP texture. Then they split the image into overlapping blocks. In the feature extraction step, they used the Gray Level Co-occurrence Matrix (GLCM) to extract features from all blocks. Finally, in the feature matching step, they used the improved kd tree algorithm to find similar blocks.

In another study, Kumar *et al*. investigated a block-based method with a hybrid of DCT and SVD algorithms [17]. They converted the RGB image into a gray image in the pre-processing step, applied the SWT algorithm on the gray image, and divided the low level (LL) band image obtained from SWT into overlapped blocks. Then in the feature extraction step, they used DCT and SVD algorithms to extract reduced feature vectors from all blocks. Finally, in the feature matching step, they used Euclidean distance to find similar blocks.

Unlike the block-based method, the key-point-based method extracts feature points from the high entropy regions and describes local features without dividing the image. Therefore, the key-point-based methods are faster than the block-based methods. Alberry *et al*. proposed a key-point-based method for copy-move forgery detection [18]. They used the SIFT algorithm to extract key points from the image in the feature extraction step. Then they used the Fuzzy C-means (FCM) clustering algorithm to match similar features of these key points in the feature-matching step.

Another study [19] proposed a key-point-based method that used a combination of two key-point extraction algorithms. They extracted key-points with SURF and A-KAZE algorithms in the feature extraction step to obtain sufficient key points. Then they used Euclidean distance to evaluate the similarity of two key-point descriptors. They sorted these distances and used the 2NN algorithm to detect similar key-points.

Most of the existing copy-move forgery methods fail to detect forgery in smooth areas. To solve this issue, Fatima *et al*. presented a two-step key-point-based forgery detection method [20]. They used the SIFT algorithm to detect keypoints in smooth regions and the FAST descriptors to detect key-points from missing regions in the feature extraction step. Then in the feature matching step, they matched key-points using the generalized 2nd nearest neighbor algorithm.

Some researchers use a hybrid of block-based and key-point-based methods to exploit the advantages of both methods. For instance, Sreelakshmy proposed a hybrid method for forgery detection [21]. They split the image into square overlapping blocks and extracted key-points using the SURF algorithm from all blocks. Then they compared blocks based on their key-points and found similar blocks if the number of similar key points exceeded a preset threshold.

Another study investigated a hybrid method that used DCT and Oriented FAST and Rotated BRIEF (ORB) algorithms [22]. They converted the RGB image into a gray image in the pre-processing step and divided the gray image into overlapping square blocks. In the feature extraction step, they used the ORB algorithm and the DCT algorithm to extract key-points and feature vectors from all blocks, respectively. Then in the feature matching step, they matched the extracted DCT features based on Euclidean distance and the extracted ORB key-points using the k-NN algorithm based on Hamming distances to detect similar blocks. Table 1 summarizes the reviewed studies.

Table 1. Summary of copy-move forgery detections.

Year	Method	Summary
2018 [15]	Block-based	Feature extraction: (DCT + PCA) and Feature matching: (Euclidean distance)
2016 [16]	Block-based	Feature extraction: (CoLBP + GLCM) and Feature matching: (Kd-tree)
2023 [17]	Block-based	Feature extraction: (DCT + SVD) and Feature matching: (Euclidean distance)
2017 [18]	Key-point-based	Feature extraction: (SIFT) and Feature matching: (Fuzzy C-mean clustering)
2018 [19]	Key-point-based	Feature extraction: (Surf + A-KAZE) and Feature matching: (2 Nearest Neighbor)
2022 [20]	Key-point-based	Feature extraction: (SIFT+FAST) and Feature matching: (Generalized 2nd nearest neighbor)
2019 [21]	Hybrid	Feature extraction: (SURF) and Feature matching: (Comparing the number of similar key points)
2020 [22]	Hybrid	Feature extraction: (DCT + ORB) and Feature matching: (Euclidean distance, K Nearest Neighbor based on Hamming distance)

3. Background

The aim of this study is to detect copy-move forgery. To achieve this, we use a combination of two meta-heuristic algorithms: the genetic algorithm or the artificial bee colony algorithm,

which identify the suspected regions of forgery, and the simulated annealing algorithm, which refine the detection of these regions. In this section, we provide a brief overview of these three algorithms.

3.1. Genetic algorithm (GA)

The genetic algorithm is a stochastic optimization algorithm that is inspired by Darwin's theory of natural selection. This algorithm is a population-based search algorithm that can solve optimization problems with complex and unknown search spaces. The main components of the GA are the solution representation, the selection, crossover, and mutation operators, and the fitness function evaluation [23-25]. The following steps briefly describe the genetic algorithm:

1- The parameters of the algorithm such as population size, maximum number of iterations, and fitness function are specified according to the problem.
2- An initial population of candidate solutions is randomly generated and the iteration index is set to zero.
3- The fitness value of each solution is calculated.
4- A subset of the current solutions is selected based on their fitness values. The cross-over and mutation operators are applied to these selected solutions to produce new solutions.
5- The old population is replaced by the new population and the iteration index is incremented.
6- If the iteration index reaches the maximum iteration, the best solution of the population is returned as the final solution; otherwise, the process goes back to step 3. Figure 2 illustrates the steps of the genetic algorithm.

3.2. Simulating Annealing (SA) algorithm

Simulated annealing (SA) is another optimization algorithm that simulates the annealing process of metals. SA iterates according to a variable temperature parameter that mimics the cooling of the metals. This algorithm starts with a high temperature and gradually decreases it to approach the optimal solution [26]. The steps of the SA algorithm are briefly stated as follows:

1- The algorithm parameters such as initial temperature, cooling function, termination condition, and fitness function are set.
2- A random initial solution is generated based on the problem and its fitness value is calculated.
3- New solutions are created around the current solution by applying suitable operations.
4- The fitness value of each new solution is computed. If this value is better than the fitness value of the current solution, the new solution is replaced; otherwise, the new solution is replaced using the probability factor with the Metropolis rule. Eq. (1) introduces this rule.

$$P = \begin{cases} 1, & \text{if } F(X_{new}) < F(X_{old}) \\ \exp\left(-\frac{F(X_{new}) - F(X_{old})}{T}\right) & \text{if } F(X_{new}) >= F(X_{old}) \end{cases} \quad (1)$$

5- The temperature value is reduced by the cooling function.
6- The algorithm terminates if the termination condition is met; otherwise, it goes back to step 3. The steps of the SA algorithm are illustrated in Figure 3.

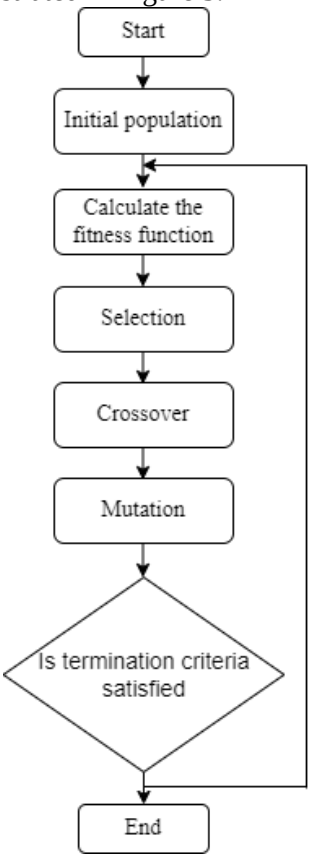

Figure 2. Flowchart of genetic algorithm.

3.3. Artificial bee colony (ABC) algorithm

The artificial bee colony (ABC) algorithm is another optimization algorithm that mimics the behavior of bees in finding food sources. In this algorithm, a food source's position represents a potential solution to the optimization problem and its nectar amount corresponds to the fitness function of the related solution. This algorithm aims to find the food source with the most nectar.

The steps of the ABC algorithm using three types of bees, employed, onlooker, and scout bees, are given below [27, 28]:

1- The algorithm parameters such as the number of food sources, employed, onlooker, and scout bees, maximum number of iterations, and fitness function, are set.
2- A set of candidate solutions is created as food sources based on the problem.
3- Employed bees visit food sources and measure the nectar amount in each source.
4- After assessing food sources, the onlooker bees choose food sources based on their nectar amounts.
5- Scout bees explore areas to find new food sources.
6- The best-found food source is remembered.
7- The algorithm terminates if the termination condition is met; otherwise, it goes back to step 3. The steps of the ABC algorithm are illustrated in Figure 4.

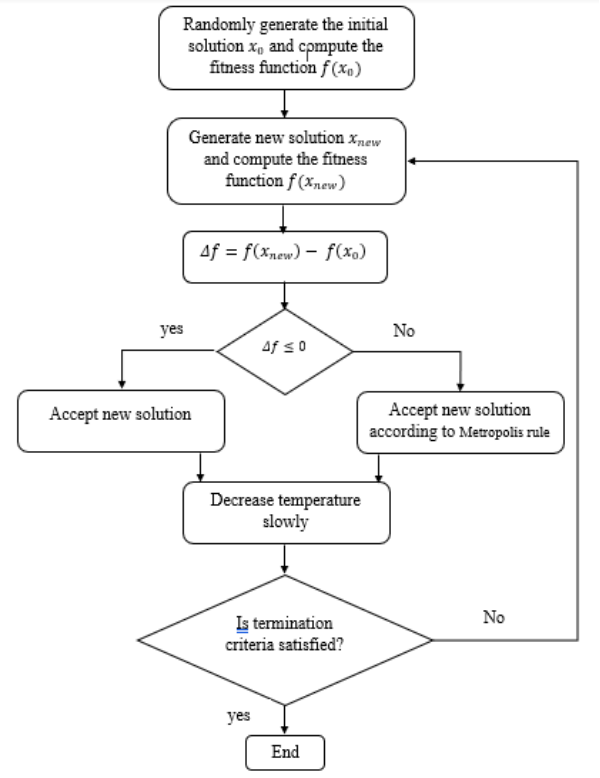

Figure 2. Flowchart of SA.

4. Proposed Method

Figure 5 summarizes the proposed method. The following sections explain each step of the method in detail.

4.1. Pre-processing

Some forgery detection algorithms start with pre-processing. This step involves operations such as converting color images to grayscale, resizing images, and equalizing image contrast.

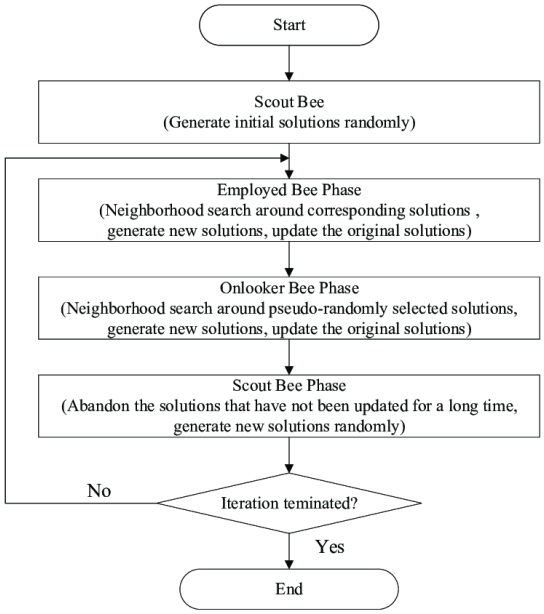

Figure 3. Flowchart of ABC algorithm.

4.2. Finding suspected blocks of forgery using genetic/ABC algorithm

The next step is to find suspected forgery blocks using a combination of block-based and key point-based methods. The basic block-based method requires comparing all pairs of blocks, which is very time-consuming and complex. To overcome this problem, we use the genetic algorithm or artificial bee colony algorithm to compare only a few block pairs based on their number of key points. This way, we can identify suspected forgery block pairs. The following sections describe how to use the genetic algorithm to find these block pairs in detail.

4.2.1. Initial population of genetic /ABC algorithm

All meta-heuristic algorithms start with generating initial and random solutions. In the genetic algorithm, each solution is a chromosome, and each chromosome has a number of genes. The number of chromosomes and genes depends on the problem. In this study, we randomly generate 100 chromosomes with 6 genes as the initial population. The first 4 genes of each chromosome represent the x and y coordinates of the left and top corner points of the block pairs, and the last 2 genes indicate the height and width of the blocks. Figure 6 shows a chromosome example.

It should be noted that the proposed chromosome is not suitable for detecting multiple forgery regions. To solve this problem, the genes of the chromosome should be modified in such a way that instead of discovering the x and y coordinates of

two blocks suspected of forgery, it searches for the x and y coordinates of multiple similar blocks.

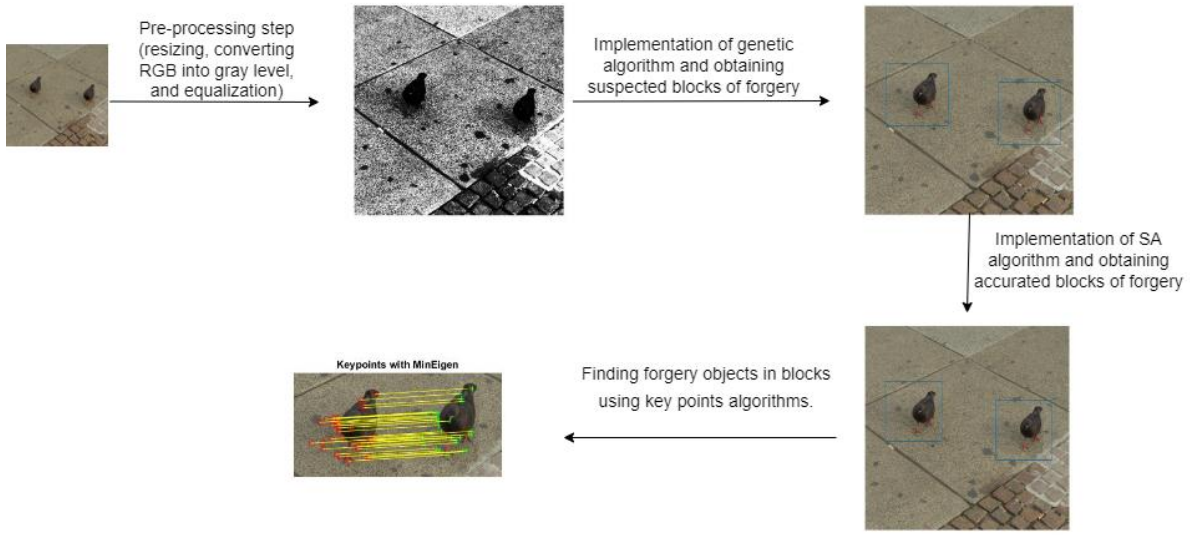

Figure 4. Flowchart of the proposed method.

X-block1	Y-block1	X-block2	Y-block2	Height	Width

Figure 5. An example of one chromosome.

The chromosomes have x and y values that are randomly generated within the range of the image rows and columns, and the width and height values of the blocks are user-defined. The width and height values of the blocks are usually large and fixed in the genetic algorithm, so that it can search for large suspected forgery blocks. Figure 7 shows three random solutions from the initial population on the selected image.

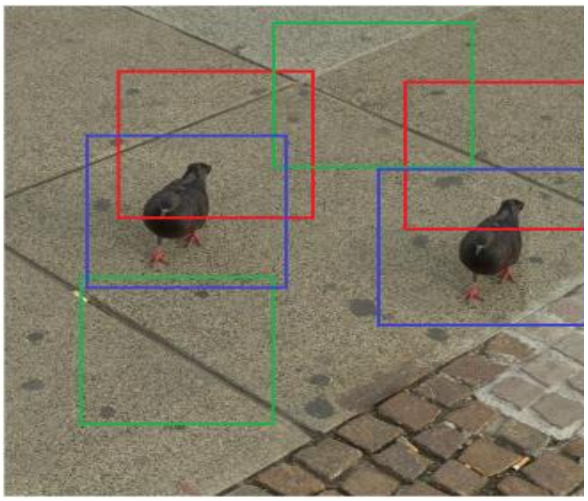

Figure 6. Three random solutions of the initial population on the selected image.

4.2.2. Fitness function of genetic /ABC algorithm

The fitness function (FF) is a crucial part of the meta-heuristic algorithm that takes a lot of time to execute. This function measures the suitability of each solution. In each iteration of meta-heuristic algorithms, the fitness function evaluates the population and selects the best solutions for the next iteration to create a new population. In this study, the fitness function is the inverse of the number of matching key points of block pairs. Eq. (2) shows this fitness function.

$$FF = \frac{1}{\text{Total number of matched keypoints of blocks}} \quad (2)$$

There are different key point extraction algorithms, such as FAST, SURF, BRISK, etc. A problem in key point-based forgery detection methods is the lack of enough key points in smooth and small regions. To address this issue, we use the total extracted key points from five methods: SURF, FAST, BRISK, Harris, and MinEigen. Figure 8 shows the extracted key points by different algorithms on suspected forgery blocks.

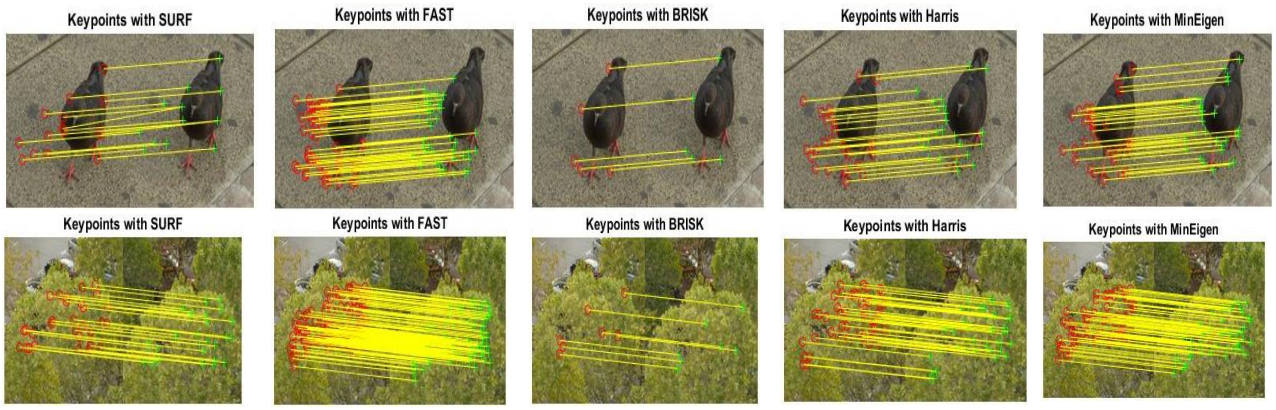

Figure 7. Extracted key points with different algorithms.

4.2.3. One-point cross-over operator of genetic algorithm

The cross-over operator generates new chromosomes in the genetic algorithm. It selects two chromosomes based on the cross-over rate and splits them from the same location into two parts. Then it combines the right part of one chromosome with the left part of another chromosome to create two new chromosomes. This way, the coordinates of the block corners are swapped and two new solutions are produced. Figure 9 shows a cross-over operator example at the third location.

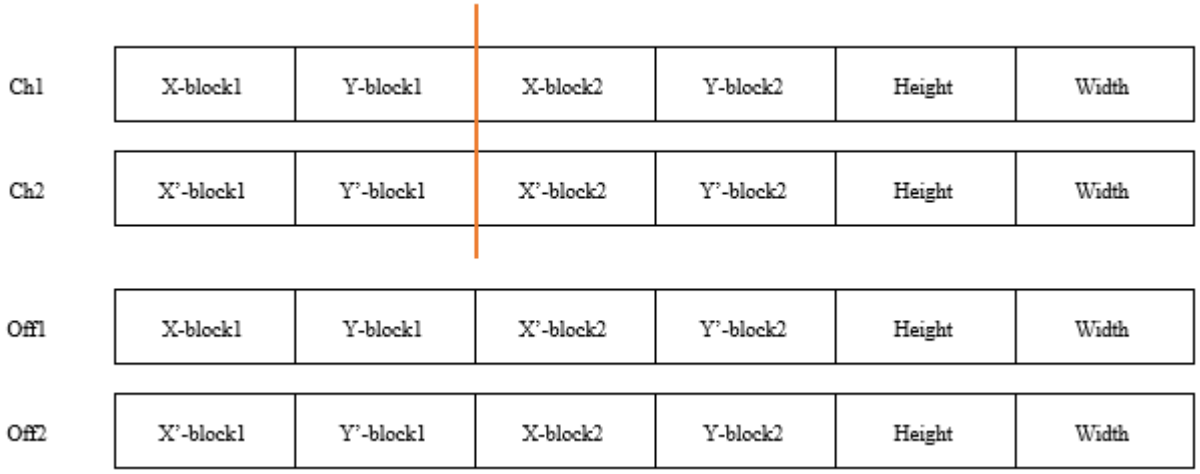

Figure 8. An example of a cross-over operator.

4.2.4. Mutation operator of genetic algorithm

Mutation is another operator in genetic algorithms. It randomly changes genes based on the mutation rate. The mutation is a useful operator; it brings back the removed genes and adds new genes to the population. In the proposed method, the mutation operator randomly changes the coordinates of one block. Figure 10 shows a mutation operator example. It randomly selects the first block and generates its x and y coordinates randomly within the range of the image rows and columns.

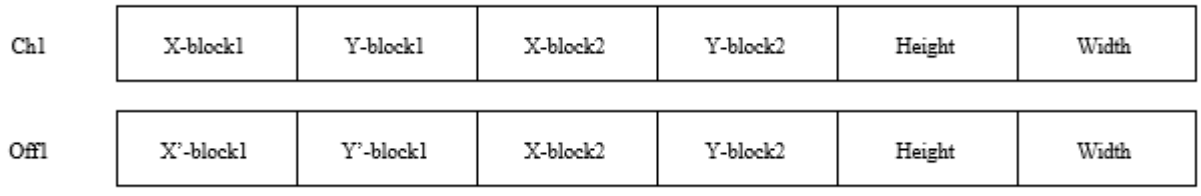

Figure 9. An example of a mutation operator.

4.2.5. Selection operator of the genetic algorithm

The selection is based on the fitness function of chromosomes in the genetic algorithm. A chromosome with a high fitness function has a higher chance of being selected. In the proposed method, each iteration applies the cross-over and mutation operators to 80% and 5% of the current population, respectively, and generates new solutions. Then it calculates the fitness functions of

new solutions and selects the next iteration population from the current population and new solutions. These processes repeat until it finds the optimal or near-optimal solution.

4.2.6. Termination condition of genetic algorithm/ABC algorithm/SA algorithm

The termination condition is a crucial and basic component of the meta-heuristic algorithms. It can be chosen from one of these options:

1- The algorithm stops after a fixed number of iterations.
2- The algorithm stops after a certain time.
3- The population does not change after several iterations.
4- The algorithm finds a solution with a pre-defined fitness function.

In this study, we use a fixed number of iterations.

4.3. Finding accurate blocks of forgery using SA

The genetic algorithm/ABC algorithm finds large blocks suspected of forgery, and then the SA algorithm refines the forgery blocks. Next, we explain the steps of the SA algorithm in the proposed method.

4.3.1. Initial solution of SA algorithm

The genetic algorithm/ABC algorithm sorts all solutions of the final iteration by their fitness functions. Then it chooses the solution with the best fitness function as the initial solution for the SA algorithm. Figure 11 shows the initial solution of the SA algorithm from the genetic algorithm/ABC algorithm.

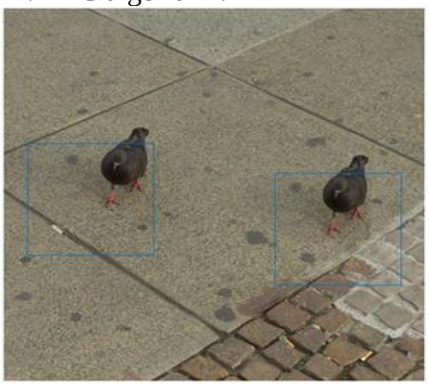

Figure 10. An example of an initial solution of SA algorithm.

4.3.2. Generate new solutions in the neighborhood of the initial solution

This step generates some new solutions around the initial solution by using suitable actions. Then it calculates the fitness function value of each new solution. If this value is better than the current solution, it accepts the new solution as the current solution; otherwise, it accepts the new solution based on the Metropolis rule probability. Eq. (3) defines the fitness function of the SA.

$$FF = \frac{1}{\text{The matched points of blocks}} - \frac{1}{\text{Height} \times \text{Width}} \quad (3)$$

The temperature is a key component of the SA algorithm. It decreases at the end of each iteration of the SA algorithm. The temperature is high in the initial iterations. High temperature increases the probability of accepting a new solution. By lowering the temperature in the final iterations, the probability of accepting new solutions decreases and the algorithm converges to the optimal or near-optimal solution. In the proposed method, the SA generates 24 new solutions by adding or subtracting a fixed value to the coordinates, width, and height of the block pair. We consider adding and subtracting zero, one, and two values. Next, we present the parameters of two algorithms, genetic and SA, in Tables 2 and 3, respectively. We selected these parameters using the cross-validation technique.

Table 2. Parameters of genetic algorithm.

Population size	Maximum iteration	Crossover rate	Mutation rate
100	100	0.80%	0.05%

Table 3. Parameters of SA algorithm.

Maximum iteration	Number of neighboring solutions	Initial temperature
100	24	1000

5. Expriment results

We implemented the proposed method with Matlab language and used built-in functions of the Matlab platform for different key point extraction algorithms.

5.1. Dataset

We used two forgery datasets, COVERAGE [29] and CoMoFod [30], for evaluation. The COVERAGE dataset has 100 forgery images with similar but genuine objects, which make forgery more realistic and challenging. These images were created by the copy-move method with different geometric transformations like scaling and rotation. This dataset has no post-processing operations. Figure 12 shows some examples of this dataset.

The CoMoFod dataset contains 5000 forged images with 512 × 512 pixels. It used the copy-move method to create forged images. The forged images in this dataset have geometric transformations such as scaling and rotation, and different post-processing operations such as JPEG Compression (JC) with different quality factors, Image Blurring (IB), Noise Adding (NA) using the

average filter with different sizes ([3 × 3, 5 × 5, and 7 × 7]), Brightness Change (BC), Color Reduction (CR), and Contrast Adjustments (CA) that are not present in other datasets. Figure 13 shows some examples of this dataset.

Figure 11. Some examples of the COVERAGE dataset. First row: Forgery images, second row: Ground truth images.

Figure 12. Some examples of the CoMoFod dataset. First row: Forgery images, second row: Ground truth images.

5.2. Evaluation metrics

Forgery detection is investigated in two levels: image forgery detection and pixel forgery detection. These two levels are evaluated with some of the standards evaluation metrics such as Accuracy (ACC), Recall (R), Precision (P), F1 score, and Intersection over Union (IoU) [31]. These metrics have been obtained from four components of the confusion matrix: True Positive (TP), False Positive (FP), False Negative (FN), and True Negative (TN). Table 4 defines the evaluation metrics.

5.3. Main result

We evaluated our method on the images of two datasets: COVERAGE and the images with different post-processing operations such as blurring, noise adding using the average filter with different sizes ([3 × 3, 5 × 5]), and JPEG compression with different quality factors ([20, 30,

80, and 90]) from CoMoFod. We first used the genetic algorithm with 100 initial random solutions to detect the suspected forgery blocks. Then we used the best solution of the genetic algorithm as the starting point of the SA algorithm. This algorithm moved to the optimal or near-optimal solution by defining a suitable operation in the neighborhood of the initial solution. Finally, we detected the forged objects in pair blocks by using key-point algorithms. Figure 14 shows some results of our method on some images.

Table 4. Different evaluation metrics.

Evaluation metrics	Formula
Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$
Precision (P)	$\frac{TP}{TP+FP}$
Recall (R)	$\frac{TP}{TP+FN}$
F1	$\frac{2\times R\times P}{R+P}$
IoU	$\frac{TP}{TP+FN+FP}$

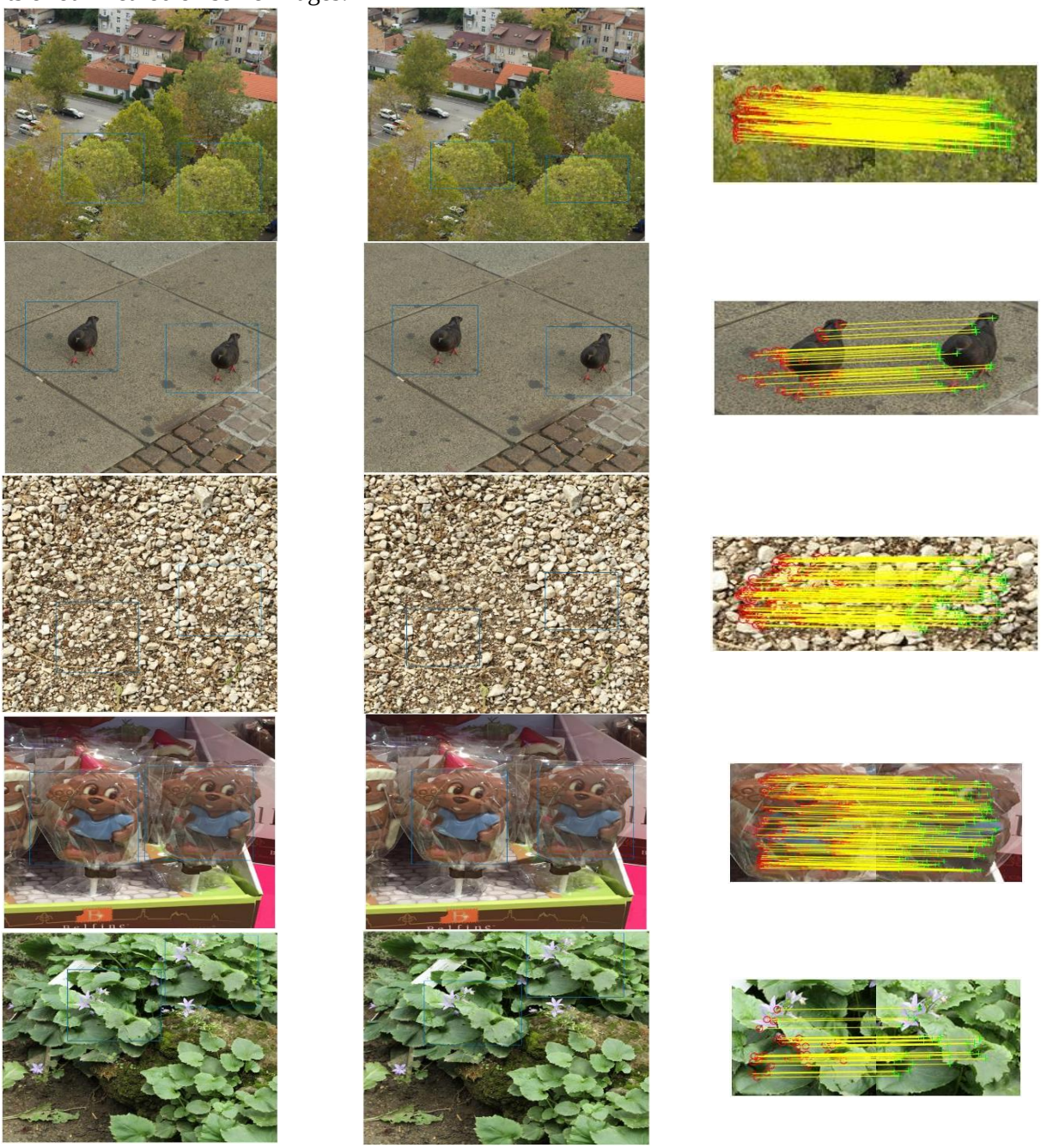

Figure 13. Result of the proposed method on some images. First column: Suspected forgery blocks obtained from the genetic algorithm, the second column: Accurate forgery blocks obtained from the SA algorithm, third column: Matched key points of forgery objects.

We should mention that the proposed method performs well in detecting forgery regions with geometric transformations and post-processing operations.

Because in the proposed method, we used a combination of five keypoint extraction algorithms: SURF, FAST, BRISK, Harris, and MinEngin. Figure 15 shows the proposed method result on some forgery images with geometric transformations and post-processing operations.

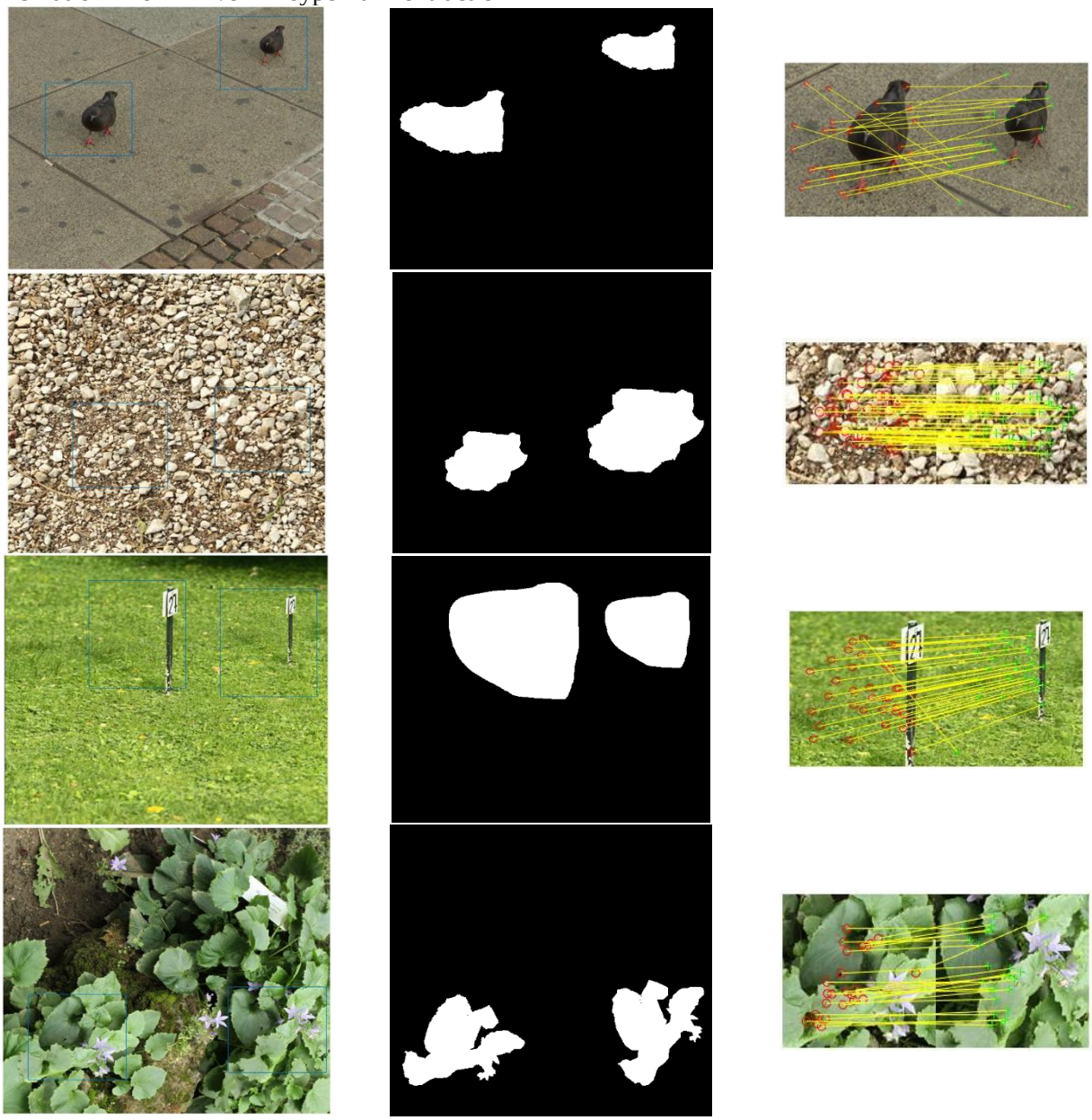

Figure 14. Result of the proposed method on some images with geometric transformations and post-processing operations. First column: Forgery blocks obtained from the SA algorithm, second column: Ground truth images, and third column: Matched key points of forgery objects.

5.4. Comparison and discussion

Forgery detection can be performed at two levels: image level and pixel level. Some studies only detect the forged image, while others detect both the forged image and the forged pixels. There are two traditional methods for copy-move forgery detection: block-based and key-point-based. Studies show that block-based methods have high computational time and complexity, and they are not robust to some geometric transformations and post-processing. Key-point-based methods are more robust to geometric transformations, but they are not effective in detecting small forgery regions due to the lack of sufficient key-points. Both block-based and key-point-based methods have different steps of pre-processing, feature extraction, and feature matching. The parameters of these steps must be adjusted individually according to the images of the dataset to detect forgery. Therefore, each forgery detection method has its own

parameters and settings, and it has been evaluated on specific forgery types and datasets. When the type of forgery or the dataset changes, the method may lose its effectiveness. In following, a comparison of the proposed method with a block-based method, a key-point-based method, and a hybrid method on the same images from the COVERAGE and CoMoFod datasets is given in Table 5.

Table 5. Comparison of the proposed method with other methods.

Method	ACC	P	R	F1	IoU	Times (s)
Block-based method (DCT) [32]	95	99.21	82.26	89.94	75.73	192.28 s
Key point-based method (DWT + SIFT) [33]	90.11	75.71	100	86.17	86.93	51.12 s
Hybrid method (DCT+ORB) [22]	85.5	84.31	87	85.58	84.82	105.81 s
proposed method with ABC and SA algorithms (DCT)	95.47	90.80	89.34	90.06	79.44	70.37 s
proposed method with GA and SA algorithms (LBP)	95.05	90.15	91.80	90.96	78.81	70.51 s
proposed method with GA and SA algorithms (key-point (SURF))	94.47	91.72	92.58	91.34	91.15	71.65 s
proposed method with GA and SA algorithms (5 key-point algorithms)	96.87	92.15	95.34	93.71	93.45	90.19 s

In addition to comparing the proposed method with other studies, we investigated the proposed method with different fitness functions. The experiment results show that the block-based method takes a lot of time and performs poorly against geometric transformations like rotation and scaling. The key-point-based method is robust to geometric transformations and post-processing operations, but it performs poorly in detecting small and smooth forgery regions.

According to the studies, it can be understood that forgery detection is a very challenging problem and it is still an open research topic. Some of the challenges in this field are: forgery detection in smooth and small regions, forgery detection with geometric transformations and various post-processing operations, multiple forgery detection, forgery detection using a combination of traditional and deep learning methods, and generalizing the forgery detection method on various forgery datasets.

6. Conclusion

Copy-move forgery is one of the simplest image manipulation techniques. In this paper, we proposed a hybrid method that combines block-based and keypoint-based methods. We first used the genetic algorithm to compare a limited number of pair blocks based on the number of matched keypoints and identify the suspected forgery blocks. This algorithm avoided comparing all pair blocks, so it improved the speed. Then we used the simulating annealing (SA) algorithm to find the accurate forgery blocks by generating new solutions around the best solution from the genetic algorithm. Finally, we identified the forged objects in the pair blocks using matched keypoints. We evaluated our method on images from the CoMoFod and COVERAGE datasets and we obtained the results of accuracy, precision, recall and IoU with values of 96.87, 92.15, 95.34, and 93.45, respectively. The experimental results showed that our approach is almost fast and robust to geometric transformations such as rotation, scaling, and their combination, and post-processing operations such as blurring, adding noise, and JPEG compression.

However, our method is not effective in detecting very small and smooth forgery regions and multiple forgery regions. In future work, we plan to use deep learning methods to solve the first problem and compare multiple blocks in each chromosome instead of two blocks in meta-heuristic algorithms to solve the second problem.

References

[1] F. Z. Mehrjardi, A. M. Latif, M. S. Zarchi, and R. Sheikhpour, "A survey on deep learning-based image forgery detection," *Pattern Recognition*, Vol. 144, pp. 1-31, 2023.

[2] Z. J. Barad and M. M. Goswami, "Image forgery detection using deep learning: A survey," *In 6th international conference on advanced computing and communication systems (ICACCS)*, 2020, pp. 571-576.

[3] A. H. Saber, M. A. Khan, and B. G. Mejbel, "A survey on image forgery detection using different forensic approaches," *Adv Sci Technol Eng Syst J*, Vol. 5, No. 3, pp. 361-370, 2020.

[4] F. Z. Mehrjardi, A. M. Latif, and M. S. Zarchi, "Copy-move forgery detection and localization using deep-learning," *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 37, No. 9, pp. 1-21, 2023.

[5] I. A. Zedan, M. M. Soliman, K. M. Elsayed, and H. M. Onsi, "Copy Move Forgery Detection Techniques: A Comprehensive Survey of Challenges and Future Directions," *International Journal of Advanced Computer Science and Applications*, Vol. 12, No. 7, 2021.

[6] F. H. Pugar, S. Muzahidin, and A. M. Arymurthy, "Copy-Move Forgery Detection using SWT-DCT and Four Square Mean Features," *In International*

Conference on Electrical Engineering and Informatics (ICEEI), 2019, pp. 63-68.

[7] M. H. Alkawaz, G. Sulong, T. Saba, and A. Rehman, "Detection of copy-move image forgery based on discrete cosine transform," *Neural Computing and Applications*, Vol. 30, pp. 183-192, 2018.

[8] Y. Wang, L. Tian, and C. Li, "LBP-SVD based copy move forgery detection algorithm," *In IEEE international symposium on Multimedia (ISM)*, 2017, pp. 553-556.

[9] K. M. Hosny, H. M. Hamza, and N. A. Lashin, "Copy-move forgery detection of duplicated objects using accurate PCET moments and morphological operators," *The Imaging Science Journal*, Vol. 66, No. 6, pp. 330-345, 2018.

[10] T. Mahmood, T. Nawaz, R. Ashraf, M. Shah, Z. Khan, A. Irtaza, and Z. Mehmood, "A survey on block-based copy move image forgery detection techniques," *In International Conference on Emerging Technologies (ICET)*, 2015, pp. 1-6.

[11] C. C. Chen, W. Y. Lu, and C. H. Chou, "Rotational copy-move forgery detection using SIFT and region growing strategies," *Multimedia Tools and Applications*, vol. 78, pp. 18293-18308, 2019.

[12] A. Badr, A. Youssif, and M. A. Wafi, "A robust copy-move forgery detection in digital image forensics using SURF," *In 8th International Symposium on Digital Forensics and Security (ISDFS)*, 2020, pp. 1-6.

[13] M. Bilal, H. A. Habib, Z. Mehmood, T. Saba, and M. Rashid, "Single and multiple copy–move forgery detection and localization in digital images based on the sparsely encoded distinctive features and DBSCAN clustering," *Arabian Journal for Science and Engineering*, Vol. 45, pp. 2975-2992, 2020.

[14] C. Lin, W. Lu, X. Huang, K. Liu, W. Sun, and H. Lin, "Region duplication detection based on hybrid feature and evaluative clustering," *Multimedia Tools and Applications*, Vol. 78, pp. 20739-20763, 2019.

[15] T. Mahmood, T. Nawaz, A. Irtaza, R. Ashraf, M. Shah, and M. T. Mahmood, "Copy-move forgery detection technique for forensic analysis in digital images," *Mathematical Problems in Engineering*, 2016.

[16] Y. Zhu, X. J. Shen, and H. P. Chen, "Covert copy-move forgery detection based on color LBP," *Acta Automatica Sinica*, Vol. 43, N. 3, pp. 390-397, 2017.

[17] S. Kumar, S. Mukherjee, and A. K. Pal, "An improved reduced feature-based copy-move forgery detection technique," *Multimedia Tools and Applications*, Vol. 82, No. 1, pp. 1431-1456, 2023.

[18] H. A. Alberry, A. A. Hegazy, and G. I. Salama, "A fast SIFT-based method for copy move forgery detection," *Future Computing and Informatics Journal*, Vol. 3, No. 2, pp. 159-165, 2018.

[19] C. Wang, Z. Zhang, and X. Zhou, "An image copy-move forgery detection scheme based on A-KAZE and SURF features," *Symmetry*, Vol. 10, No. 12, 2018.

[20] B. Fatima, A. Ghafoor, S. S. Ali, and M. M. Riaz, "FAST, BRIEF and SIFT based image copy-move forgery detection technique," *Multimedia Tools and Applications*, Vol. 81, No. 30, pp. 43805-43819, 2022.

[21] I. J. Sreelakshmy, and B. C. Kovoor, "Hybrid Method for Copy-Move Forgery Detection in Digital Images," *In Proceedings of the International Conference on ISMAC in Computational Vision and Bio-Engineering 2018 (ISMAC-CVB)*, 2019, pp. 119-127.

[22] V. Mehta, A. K. Jaiswal, and R. Srivastava, "Copy-move image forgery detection using DCT and ORB feature set," *In Futuristic Trends in Networks and Computing Technologies: Second International Conference, FTNCT 2019, Chandigarh, India*, 2020, pp. 532-544.

[23] S. Katoch, S. S. Chauhan, and V. Kumar, "A review on genetic algorithm: past, present, and future," *Multimedia Tools and Applications*, Vol. 80, pp. 8091-8126, 2021.

[24] M. A. Albadr, S. Tiun, M. Ayob, and F. Al-Dhief, "Genetic algorithm based on natural selection theory for optimization problems," *Symmetry*, Vol. 12, No. 11, 2020.

[25] Z. Mehrnahad, A. M. Latif, J. Z. Ahmadabadi, "A New Scheme for Lossless Meaningful Visual Secret Sharing by using XOR Properties," *Journal of AI and Data Mining*, Vol. 11, No. 2, pp. 195-211, 2023.

[26] A. H. Zhou, L. P. Zhu, B. Hu, S. Deng, Y. Song, H. Qiu, and S. Pan, "Traveling-salesman-problem algorithm based on simulated annealing and gene-expression programming," *Information*, Vol. 10, No. 1, 2018.

[27] D. Karaboga and B. A. Akay, "Comparative study of artificial bee colony algorithm," *Applied mathematics and computation*, Vol. 214, No. 1, pp. 108- 132, 2009.

[28] F. N. A. Baharudin, N. A. Ab. Aziz, M. R. Abdul Malek, A. K. Ghazali, and Z. Ibrahim, "Indoor Comfort and Energy Consumption Optimization Using an Inertia Weight Artificial Bee Colony Algorithm," *Algorithms*, Vol. 15, No. 11, 2022.

[29] B. Wen, Y. Zhu, R. Subramanian, T. T. Ng, X. Shen, and S. Winkler, "COVERAGE—A novel database for copy-move forgery detection," *In IEEE international conference on image processing (ICIP)*, 2016, pp. 161- 165.

[30] D. Tralic, I. Zupancic, S. Grgic, and M. Grgic, "CoMoFoD—New database for copy-move forgery detection," *In Proceedings ELMAR-2013*, 2013, pp. 49-54.

[31] O. M. Al-Qershi and B. E. Khoo, "Evaluation of copy-move forgery detection: datasets and evaluation

metrics," *Multimedia Tools and Applications*, Vol. 77, pp. 31807-31833, 2018.

[32] T. Mahmood, T. Nawaz, A. Irtaza, R. Ashraf, M. Shah, and M. T. Mahmood, "Copy-move forgery detection technique for forensic analysis in digital images," *Mathematical Problems in Engineering*, 2016.

[33] M. F. Hashmi, A. R. Hambarde, and A. G. Keskar, "Copy move forgery detection using DWT and SIFT features," *In 13th International conference on intellient systems design and applications*, 2013, 188-193.

یک روش ترکیبی بهینه برای تشخیص جعل کپی-انتقال

فاطمه زارع مهرجردی، علی محمد لطیف[1,*] و محسن سرداری زارچی[2]

[1] دانشکده مهندسی کامپیوتر، دانشگاه یزد، یزد، ایران.

[2] دانشکده مهندسی کامپیوتر، دانشگاه میبد، میبد، یزد، ایران.

ارسال ۲۰۲۳/۰۵/۲۵؛ بازنگری ۲۰۲۳/۰۷/۲۰؛ پذیرش ۲۰۲۳/۰۸/۱۹

چکیده:

تصویر یک ابزار ارتباطی قدرتمند است که به‌طور گسترده در کاربردهای مختلف مانند پزشکی قانونی و دادگاه استفاده می‌شود، جایی که اعتبار تصویر بسیار اهمیت دارد. با توسعه و در دسترس بودن ابزارهای ویرایش تصویر، دست‌کاری تصویر به راحتی و برای اهداف خاص قابل انجام است. جعل کپی-انتقال یکی از ساده‌ترین و رایج‌ترین روش‌های دست‌کاری تصویر است. دو روش سنتی برای تشخیص این نوع جعل وجود دارد: روش مبتنی بر بلوک‌بندی و روش نقاط کلیدی. در این مطالعه، ما یک رویکرد ترکیبی از روش‌های مبتنی بر بلوک‌بندی و نقاط کلیدی با استفاده از الگوریتم‌های فراابتکاری برای یافتن روشی بهینه ارائه کردیم. برای این منظور، ابتدا جفت بلوک‌های مشکوک به جعل با استفاده از الگوریتم ژنتیک و با حداکثر تعداد نقاط کلیدی منطبق به عنوان تابع هزینه جستجو شده است. سپس با استفاده از الگوریتم شبیه‌سازی تبرید و تولید راه‌حل‌های همسایه در اطراف بلوک‌های مشکوک، بلوک‌های دقیق یافته شده است. روش پیشنهادی بر روی مجموعه داده‌های CoMoFod و COVERAGE ارزیابی شده و نتایج دقت، صحت، فراخوان و IOU به ترتیب با مقادیر ۹۶,۸۷، ۹۲,۱۵، ۹۵,۳۴ و ۹۳,۴۵ به دست آمده است. نتایج ارزیابی عملکرد رضایت‌بخش روش پیشنهادی را نشان می‌دهد.

کلمات کلیدی: روش بلوک‌بندی، روش نقاط کلیدی، روش ترکیبی، الگوریتم ژنتیک، الگوریتم شبیه‌سازی تبرید.