IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 3 April 2024, accepted 3 May 2024, date of publication 6 May 2024, date of current version 14 May 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3397466

RESEARCH ARTICLE

Advancing Copy-Move Manipulation Detection in Complex Image Scenarios Through Multiscale Detector

ANJALI DIWAN[1], (Senior Member, IEEE), RAJESH MAHADEVA[2], (Member, IEEE), AND VINAY GUPTA[2], (Member, IEEE)

[1]Department of CE-AI, Marwadi University, Rajkot, Gujarat 360003, India
[2]Department of Physics, Khalifa University, Abu Dhabi, United Arab Emirates

Corresponding authors: Vinay Gupta (vinay.gupta@ku.ac.ae) and Anjali Diwan (anjali.diwan@ieee.org)

ABSTRACT This research presents a new approach for identifying instances of copy-move forgeries in digital images by utilizing the Multiscale Detector a Neural Network-based method, which serves as an image key-point detector and descriptor. The act of copy-move manipulation involves the replication and subsequent insertion of a specific segment of an image, intending to modify the overall content of the image. The approach we utilize leverages the sophisticated functionalities of Multiscale Detector, a framework that combines key-point detection with descriptor extraction, to accurately detect and localize instances of copy-move manipulation. The effectiveness of our approach is assessed on a range of copy-move forgeries, encompassing instances that have undergone post-processing and geometric transformations. The experimental findings illustrate the resilience of our approach in identifying instances of manipulations over a diverse range of textured images and various alteration approaches. Furthermore, our approach demonstrates strong performance even when subjected to supplementary processing procedures such as brightness modification, color reduction, contrast adjustment, and blurring. Our suggested method has greater performance when compared to the existing manipulation detection approach, as demonstrated through a comparative analysis. In addition, the algorithm we have developed has high computing efficiency, allowing for real-time detection of forgeries. The methodology employed in this study, which builds upon the Multiscale Detector framework, offers a highly effective approach to the detection of copy-move manipulations in digital images.

INDEX TERMS Copy-move forgery, multiscale detector, neural network, computer vision, image forensics, forgery detection.

I. INTRODUCTION

The availability and user-friendliness of cutting-edge image editing software, such as Adobe Photoshop, Adobe Lightroom, GNU Image Manipulation Program (GIMP), Picasa, Paint.net, and others, have made manipulating digital images simple [1], [2]. These techniques make it easier to create altered images that might seem completely real without ever revealing their fraud. Unfortunately, the unethical use of altered digital images has spread to numerous fields, including scientific journals, newspapers, magazines, websites, visual representations of medicine, and even courtrooms [3], [4]. As a result, the discipline of digital image forensics has emerged, and researchers have created a variety of techniques to identify altered photographs.

Image splicing and copy-move have emerged as the two most popular counterfeit methods. The former comprises combining many photographs to make a false composite, while the latter entails copying and moving portions of an image to produce a fake version [5]. When certain sections of an image are copied and then pasted inside of another image, a copy-move manipulation occurs, producing modified

The associate editor coordinating the review of this manuscript and approving it for publication was Xianzhi Wang.

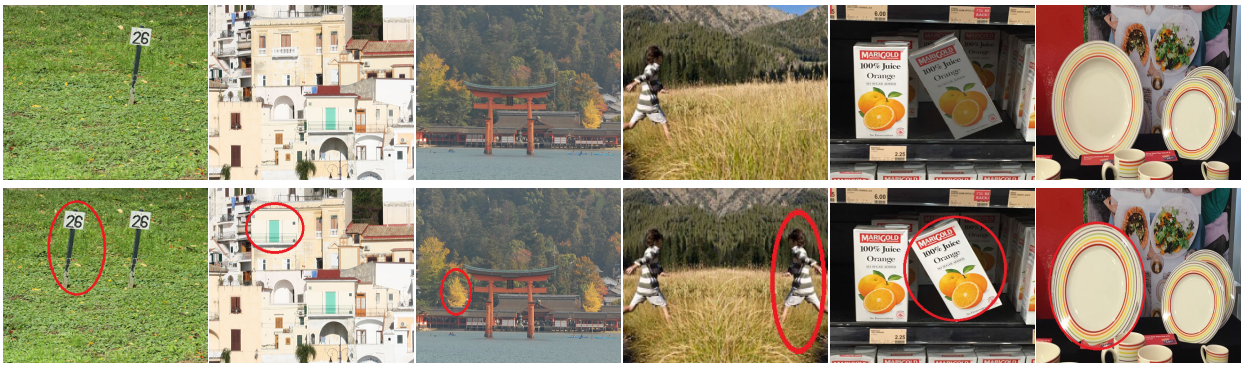

FIGURE 1. These images depict instances of copy-move manipulation. The initial row exhibits unaltered images, whilst the subsequent row showcases modified images.

content. This method can add extra aesthetic components or obscure actual information in the image, as shown in Figure (1). The detection of copy-move manipulated images is extremely difficult since the manipulated zone gets the visual traits of the surrounding areas because it is a part of the same image [6] Additionally, malevolent attacks use several techniques to conceal the modified content, including rotating, scaling, applying JPEG compression, introducing noise, and more, all of which complicate the process of manipulated image detection.

There are many instances where photographs of man-made or natural environments contain actual items with comparable looks, as shown in Figure (1). In such cases, it becomes challenging to recognize identical parts inside an image, which is a key component of copy-move manipulation detection. The difficulty in distinguishing between authentically captured content and replicated portions is caused by the presence of real but similar items in the image. This work presents a novel method for detecting copy-move manipulations to tackle this research conundrum head-on. This innovative technique boasts the ability to precisely identify the locations of altered portions inside the manipulated image in addition to detecting manipulated photographs.

A. CHALLENGES AND LIMITATIONS IN EXISTING METHODS

Copy-move manipulation stands out as a particularly nefarious type of manipulation among the deceptive techniques used. Existing detection approaches have drawbacks that limit their usefulness and efficacy. Our suggested method is motivated by addressing these limitations. The following are some of the limitations of the existing detection approaches.

- Sensitivity to Image Quality: Many current detection approaches are sensitive to changes in image quality, including noise, distortion, and compression artifacts. The accuracy of manipulation detection can be compromised by this sensitivity's potential for false positives or false negatives.
- Limited Generalisation: Existing approaches frequently have trouble adapting well to new or undiscovered types of manipulation, particularly when they must deal with a wide range of texture, color, brightness, lighting, contrast, and blur variations. Post-processing like JPEG compression, noise addition, and geometrical transformations such as angle rotation and change in scale of the manipulated region.
- Limited Image Type: There aren't many thorough studies that cover manipulated images with various attributes, such as diverse image sizes, manipulated region sizes, and image file formats, in the present corpus of research. This gap emphasizes the demand for more adaptable detection methods.
- Limited Scalability: Some approaches have difficulty handling huge datasets or require complex computations, which limits their usefulness in situations involving a large number of digital images.

B. OUR PROPOSED CONTRIBUTION

To overcome the challenges and limitations discussed above, we proposed the use of a Multiscale Detector key-point detector and descriptor. The Multiscale Detector architecture is a deep neural network-based key-point detector created for practical and real-time applications. Our proposed approach includes several significant contributions.

- Sensitivity to Image Quality: Our approach is exceptionally resilient to changes in image quality, like the presence of noise, distortion, and compression. False positives as well as false negatives have been thoroughly addressed while retaining a balanced outcome.
- Detection of Novel assaults: Our approach efficiently finds a wide range of novel attacks, such as those employing different textures, varied lighting, blur, varied colors, JPEG compression, and geometrical transformation attacks. Furthermore, it has exceptional proficiency in detecting coordinated assaults that use both rotation and scale inside the modified areas.
- Robustness Against Complex Attacks: While previous research often overlooked flip assaults and combinations of rotation, and scaling, with post-processing, our

approach excels at detecting these intricate manipulations, leading to a higher level of accuracy.
- Versatility and Robustness: Our approach shows its efficacy on a variety of datasets with images of various sizes, manipulated region sizes, and image file formats. This adaptability highlights how effective our system is in spotting copy-move manipulations in many situations.
- Processing Time Efficiency: Our approach drastically saves the processing time needed for detection, ensuring quicker outcomes without compromising accuracy.

In our research, we describe a groundbreaking approach for improving the precision, effectiveness, and adaptability of copy-move manipulation detection in digital images using key-point detection via the Multiscale Detector architecture. Our approach addresses the complex difficulties presented by diverse image transformations and is motivated by the need to get past the constraints of existing solutions.

The remaining sections of the paper are organized in the following manner: Section II comprises established methodologies. Section III comprises the Multiscale Detector and its operational mechanism. Section IV focuses on the Multiscale Descriptor network, which is used for key-point detection. Section V encompasses the methods employed in the suggested approach. Section VI encompasses the assessment criteria and dataset employed in this study. Section VII comprises the findings and interpretations of the results. Section VIII comprises the conclusion.

II. LITERATURE REVIEW

The detection of copy-move forgeries using traditional techniques can be classified into two distinct categories, which are determined by their dependence on manually designed features [7]. These categories include block-based, key-point-based, and deep learning-based approaches. Block-based approaches involve the extraction of local features by utilizing overlapping or non-overlapping patches [8], whereas key-point-based approaches focus on patches that have a high density of key-points. Deep learning-based approaches utilize deep neural networks to learn and extract features for copy-move forgery detection. All of the approaches under the passive detectionÂ category provide numerous ways to identify a copy-move manipulation in digital images. The parts that follow will provide a comprehensive analysis of these strategies.

A. BLOCK-BASED

Block-based approaches have been developed to identify instances of copy-move manipulations in images without the need for additional post-processing. However, these approaches have certain limitations when it comes to detecting geometric attacks, and they also require significant computational power [9], [10]. Numerous researchers have investigated a block-based methodology for identifying copy-move manipulation. Some of these investigations comprise; Babu and Rao [11], proposed to utilize a blend of several iterations of Local Binary Patterns (LBP) such as local ternary pattern, local phase quantization, and local Gabor binary pattern. The characteristics of BP are employed to train the classifier model, while the Support Vector Machine is utilized for the classification of copy-move manipulation verification. Diwan et al. [12] proposed a method that uses Locality Preserving Projection to detect instances of copy-move manipulation. Their block-based approach exhibits efficacy for both post-processed and unaltered images. Hosny et al. [13] proposed a method for sub-sampling images using QPCETMs, which integrates the Sobel operator to detect edges and remove small sections. However, its efficacy may be restricted when used in images with uniform or densely textured areas. Gani et al. [14] proposed approach involves applying Cellular Automata to the individual Discrete Cosine Transform (DCT) block characteristics inside the image. However, it is important to note that this method has a significantly high temporal complexity.

The block-based method is ineffective in identifying when the duplicated area undergoes geometric transformations, such as rotation, scaling, or flip. This is because the transformed area may not match precisely with any of the blocks in the image [15]. Hence, an alternative methodology is required to identify geometrically altered instances of copy-move forgeries.

B. KEYPOINT-BASED

Key-point-based approaches are employed to extract and compare distinctive key-points to identify instances of image manipulation when a block-based approach fails to detect [7]. Diwan et al. [16] In this paper, the SuperPoint detector for detecting copy-move manipulations in complex assaults is described. This approach employs a trained model with self-supervised learning and a quick detection time. Kumar and Meenpal [17] utilized a silent keypoint section approach with SIFT features along with KAZE image keypoint features for the detection of copy-move forgery. Lee et al. [18] Proposed a copy-move detection methodology that utilizes a rotation-invariant characteristic and high-frequency wavelet coefficients. This approach uses a correlation module and a reduced mask decoder module. Venugopalan and Gopakumar [19] employed SIFT keypoint with DBSCAN for the clustering of the extracted keypoints. Additionally, they used the Hu invariant moment for getting for identification of similar regions in copy-move images. Wang et al. [20] utilize simple linear iterative clustering (SLIC) and the K-multiple-means (KMM) for feature extraction along with Fast Quaternion Generic Polar Complex Exponential Transform (FQGPCET) and the texture features based on the Gray-level co-occurrence matrix (GLCM) to enhance the robust feature extraction for copy-move forgery detection.

Keypoints are often identified by locating portions of the image with high-contrast variations in texture or colour. The number of keypoints detected in an image is determined by its

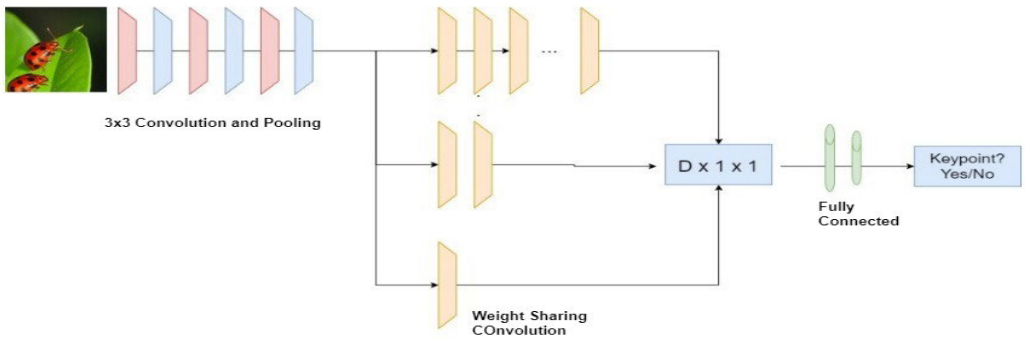

FIGURE 2. The architecture for training a network of multiscale detector.

texture. Smooth images may have fewer detectable keypoints than images with more textured regions. This may result in a decreased detection rate for keypoint-based techniques for smooth images [21].

C. DEEP LEARNING-BASED

However, it is important to acknowledge that block-based and keypoint-based approaches include certain limitations, including the requirement for robustness against various image processing techniques and potential issues related to computing efficiency [22], [23]. This phenomenon has prompted researchers to investigate deep learning frameworks as a potential choice, which have demonstrated encouraging outcomes in enhancing the precision of copy-move manipulation detection [24], [25]. Numerous methodologies utilizing deep learning frameworks have been put forth in the realm of copy-move manipulation detection. Zhu et al. [26] In this study they have proposed end-to-end AR-Net along with deep matching to capture context information fully. Babu and Rao [27] In this study they have used an optimized naive base classifier along with GLCM and steerable pyramid transform for copy-move forgery detection. Rhee [28] In this study they used novelty GT images for image classification and semantic segmentation. Mashael et al. [29] In this study they used a Neural architecture search network for feature extraction and manipulation detection, and they tuned parameters using the RSA approach. Khalil et al. [30] In this study they used transfer learning for image copy-move manipulation detection. Their results suggest that MobileNetv2 gives the best results. Rao et al. [31] In this study they used RestNet with transformer decoding for addressing splicing and copy-move manipulation detection.

III. WORKING OF MULTISCALE DETECTOR

The Multiscale Keypoint Detector [32] is a sophisticated neural network specifically created to detect key-points in digital images at different scales. This detector employs a fusion of convolutional and pooling layers, alongside recursive convolutions, to analyze image content and produce a feature map. Significantly, the network adjusts to various keypoint scales by employing a scale-dependent branching mechanism. The primary objective is to identify important areas inside images, making the Multiscale Keypoint Detector a desirable tool for tasks such as copy-move forgery detection. In this context, the ability to recognize key-points across several scales is essential for reliable and precise analysis of image integrity.

A. TRAINING ARCHITECTURE FOR MULTISCALE KEY-POINT DETECTION NETWORK

The main goal of the detection system is to accurately identify areas in an input image that can be considered high-quality key-points. The architecture of the multiscale key-point detection network, as depicted in Figure (2), involves the application of convolutional and pooling layers to image patches is subsequently followed by recursive convolution until the feature-map dimension is reduced to 1×1. Considering that a batch can contain patches of different sizes, a branch that depends on the scale is selected for each patch, which determines the required number of recursive convolutions. The final phase comprises two completely connected layers that direct the network towards a binary classification of keypoints. This architecture is designed to facilitate the efficient learning of keypoints at many scales, allowing for differences in scale across the dataset.

During the training procedure, the network takes sets of patches $\{p_i\} \subset P$ as input, accompanied by binary labels that indicate if each patch represents a good key-point. The detection network operates as a binary classification Convolutional Neural Network, learning to discern if a specific patch qualifies as a good key-point. A scale-dependent branching mechanism is employed for each patch, dynamically determining the number of recursive convolutions based on the scale. To enhance performance, the training incorporates hard-negative mining. This involves the random sampling of the dataset to construct batches with a mixture of positive and negative patches, thereby improving the key-point detector's efficacy.

Training Objective: The training objective is defined by the loss function $L_{K,P}$, which consists of two terms. The first term employs hinge loss to model key-point detection as a

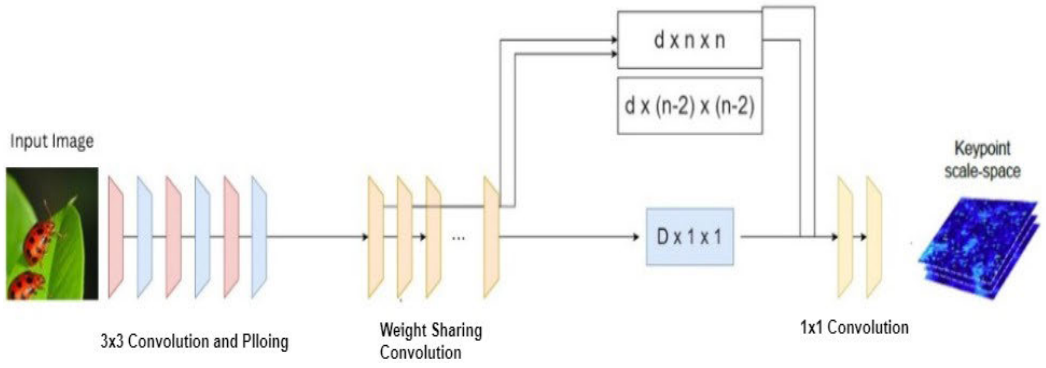

FIGURE 3. Inference architecture of multiscale detector.

binary classification task. The second term utilizes a squared difference loss, introducing a Gaussian-like reaction in the vicinity of the patch's center to discourage network reactions on patches that are not centered. The loss function is defined as the combination of multiple factors.

$$L_{KP} = \frac{1}{N} \sum_{j} \left(\lambda \max(0, 1 - y_j x_j) + (1 - \lambda) \left\| x_j - h_j \right\|^2 \right) \tag{1}$$

here:
x_j: Network output
y_j: Training label ($y_j (\in \{-1, 1\})$)
h_j: Gaussian-like reaction in the vicinity of the patch's center.

B. INFERENCE ARCHITECTURE

The inference architecture of the multiscale keypoint detection network commences by subjecting the input image to a series of convolutions and pooling layers. Subsequently, a recurrent convolutional procedure is employed until the feature map size is diminished to 1×1. Following each iteration of recursive convolution, the network calculates the keypoint feature map. Significantly, as the convolutions advance further into the network, the receptive fields of individual neurons expand, leading to output feature maps that bear resemblance to a scale-space of keypoints. This design guarantees that the network efficiently gathers information at several scales, resulting in a thorough representation of important features across various scales in the input image. During inference as shown in Figure (3) the network processes whole images assuming they are at least of size 64×64. The architecture includes:

1) Convolutional and pooling layers are applied to the input image.
2) Recursive convolution is applied until the feature-map dimension is 1×1.
3) The network is fully convolutional, outputting a feature-map in which each value represents the score of a particular image region's key-point.

Inference Procedure:

1) The network outputs a feature-map after every recursive convolution, resembling a key-point scale space.
2) The generated feature-maps enable the identification of the optimal scale for each patch by determining the scale with the highest key-point value.

Training Objective: The joint loss function L_{KP}combines hinge-loss and squared difference loss terms to train the key-point detection network:

$$L_{KP} = \frac{1}{N} \sum_{j} \left(\lambda \max(0, 1 - y_j x_j) + (1 - \lambda) \| x_j - h_j \|^2 \right) \tag{2}$$

These equations represent the training and inference procedures for a multiscale key-point detection network, emphasizing the scale-dependent branching mechanism and the convolutional architecture's ability to handle key-points of varying scales efficiently.

IV. DESCRIPTOR NETWORK FOR KEY-POINT MATCHING

The Multiscale descriptor is a feature representation that effectively captures information at several levels of scale within an image. In the domain of key-point detection and matching, the multiscale descriptor is specifically crafted to exhibit resilience against fluctuations in the dimensions and visual characteristics of key-points.

A. WORKING OF MULTISCALE DESCRIPTOR

A multiscale descriptor operates by extracting significant data from image patches at various resolutions. The method usually commences by identifying key-points at different scales through a key-point detection network. After identifying the key-points, patches that are centered around these key-points are extracted from the image at various scales, creating a multiscale database. The multiscale descriptor network is subsequently trained to acquire a nonlinear feature embedding from these patches and map them into a feature space. This embedding is specifically designed to minimize the Euclidean distance between descriptors of matching

patches while maximizing the distance between descriptors of non-matching patches. During the process of inference, the multiscale descriptor network, which has been trained, analyses complete images and produces a feature-map that accurately reflects the key-points and their descriptors at various sizes. This allows the system to rapidly detect and align distinctive features at different scales, making it resistant to variations in size and perspective within the images. Utilizing multiscale descriptors improves the effectiveness of tasks that rely on key-points, such as picture registration and copy-move manipulation detection. This is achieved by offering a full representation that takes into account key-point information at various resolutions.

The objective of the descriptor network is to learn a nonlinear feature embedding $f(p)$ from image patches p into a feature space R^d. This embedding ensures that the Euclidean distance between the embeddings of two patches is minimal if they are a match and significant if they are not. The training employs a triplet network methodology, where a patch p_1 is expected to exhibit greater proximity to a positive match p_2 than to a negative patch p_3.

Training Procedure

Input: Set of patch triplets p_1, p_2, p_3.

Architecture: The embedding feature vector in every patch is computed using a convolutional neural network, with three networks utilizing identical weights. The feature vectors are scaled to conform to the d-dimensional unit hypersphere.

Euclidean Distances: Calculate the combinations between each pair of items. The Euclidean distances calculated are based on the feature vectors of the anchor patch and the positive and negative patches.

Triplet Ranking Loss: Employ a triplet ranking loss to minimize the distance between matching patches in the feature space and maximize the distance between non-matching patches. The loss is characterized as:

$$L_T = \frac{1}{N} \sum_j \max(0, D(p_{j1}, p_{j2}) - D(p_{j1}, p_{j3}) + h) \quad (3)$$

Here D is the Euclidean distance function and h is a chosen margin (0.2 in this case).

Online Hard Negative Mining

Triplet Sampling: Triplate sampling involves random selection of negative patches, while the anchor and positive patches are taken from the match set M.

Online Hard Negative Mining: To ensure convergence, sample negative patches that violate the triplet constraint the most. For each matching pair in the training batch, choose the negative patch that violates the constraint the most. Each pair of matching items in a batch can be selected from a common pool of negative patches.

Training Objective

The loss function L_T for training the key-point description network is defined by the triplet ranking loss:

$$L_T = \frac{1}{N} \sum_j \max(0, D(p_{j1}, p_{j2}) - D(p_{j1}, p_{j3}) + h) \quad (4)$$

where $D(p_a, p_b) = \|f(p_a) - f(p_b)\|_2$ represents the Euclidean distance function between the embedding feature vectors computed from image patches p_a and p_b.

This approach ensures that the descriptor network learns a nonlinear feature embedding that facilitates the accurate matching of key-points in images, effectively capturing the relationships between positive and negative patches in the training triplets. The use of online hard negative mining helps focus on challenging examples that contribute to improved learning and convergence.

The significance of the descriptor network in the field of copy-move manipulation detection is highlighted by its potential, refined through meticulous training, to provide discriminative embeddings for key-points. These embeddings function as different identifiers, encoding the individual characteristics of image patches that are crucial for efficient matching. When key-points display congruence, it provides strong evidence indicating the presence of duplicated regions within the image. This capacity to match becomes crucial in the wider challenge of detecting occurrences of copy-move forgeries in digital images. The purpose of the descriptor network goes beyond only representing features. It is a crucial component in the detection process, providing a strong mechanism to reveal manipulated areas and enhance the overall dependability and precision of manipulation detection systems.

V. METHODOLOGY

The proposed approach combines the advantages of the multiscale key-point detector and descriptor network, providing a comprehensive strategy for identifying instances of copy-move manipulation in digital images. The inclusion of multiscale information improves the system's ability to adapt to various manipulation situations, while the descriptor network guarantees precise matching, hence enhancing the overall dependability and effectiveness of the manipulation detection system.

A. FEATURE EXTRACTION USING MULTISCALE DESCRIPTOR

The Multiscale Detector is designed to learn a nonlinear function and distinguish potential manipulation regions within an image. Simultaneously, a descriptor network is employed to acquire a nonlinear feature representation from image patches. The primary objective is to ensure that the Euclidean distance between embeddings is small for matching patches and large for non-matching ones, enhancing the network's discriminative capabilities.

1) DATASET PREPARATION

Curating a thorough dataset is essential for training and assessing the copy-move manipulation detection system. The dataset should comprise a heterogeneous assortment of both genuine and manipulated images to guarantee the model's efficacy over a wide range of scenarios.

The process of dataset construction entails collecting authentic images together with manipulated ones, in which copy-move forgeries are intentionally inserted. Authentic images serve as the reference point, whereas altered images contain duplicated sections. The dataset's diversity is crucial for effectively training a resilient model that can accurately identify copy-move forgeries across different scenarios. We have used seven publicly available datasets for this purpose. Details of these datasets are given in the section

2) DESCRIPTOR LEARNING

We employ a triplet network methodology. To train a descriptor network for acquiring a nonlinear feature embedding from image patches. Each image patch undergoes convolutional processing to calculate its embedding feature vector, aiming to create embeddings that facilitate accurate matching and discrimination between patches. The multiscale key-point detector incorporates a scale-dependent branching mechanism guided by convolutional and pooling layers, adapting to varying key-point scales in the dataset.

The comprehensive training objective for both the multiscale key-point detector and the descriptor network includes hinge loss for classification and squared difference loss to penalize responses on non-centered patches. To enhance training and convergence, online hard negative mining is incorporated, focusing on challenging examples during the learning process. The applied loss functions encompass hinge loss for classification and squared difference loss, contributing to the effective training of both networks.

3) INFERENCE PHASE

In the inference phase, the multiscale key-point detection network processes entire images using convolutions and pooling layers. This fully convolutional design ensures efficient multiscale inference, accommodating the varied nature of key-points in the dataset. Triplet loss is employed during the descriptor network's training, ensuring that the embedding feature vectors of matching patches are closer than those of non-matching ones. Online hard negative mining is utilized to select challenging triplets, enhancing the learning process.

4) DETECTION PHASE

During the detection phase, the trained multiscale key-point detector identifies key-points in the input image. The descriptor network is then employed to compute embeddings for the detected key-points. Leveraging multiscale information, the matching process effectively identifies duplicated regions indicative of copy-move forgeries. This systematic approach combines the strengths of multiscale key-point detection and descriptor learning, offering a robust solution for detecting copy-move forgeries in digital images.

B. ALGORITHMIC STEPS OF DETECTION

The approach for identifying copy-move manipulation based on a Multiscale Detector is employed to identify and precisely find the regions that have been modified inside an input image. The proposed algorithm for detecting copy-move forgeries employs a multiscale key-point detection and descriptor learning approach, which functions through a series of separate stages. Initially, the digital image input is loaded for analysis. The approach then employs a multiscale key-point detection network that utilizes convolutional and pooling layers together with a scale-dependent branching mechanism. Recursive convolutions are used iteratively till the feature-map dimension is decreased to 11, resulting in a collection of key-points that represent various sizes in a shared feature space.

After identifying key-points, image patches are extracted with the key-points in their center. These patches are chosen to cover regions that are large enough to capture local information. Afterward, a network for learning descriptors is trained using triplet loss. This procedure entails passing each patch through a convolutional neural network to calculate an embedded feature vector, which is then normalized to the hypersphere of a D-dimensional unit. The network calculates the pairwise Euclidean distances between the embeddings of the anchor, positive, and negative patches. This process helps the network to project matching patches closer together and non-matching patches farther apart.

After acquiring embeddings, the technique computes pairwise distances and establishes a matching threshold to identify probable key-point matches. Identifying corresponding key-points exposes areas that suggest the presence of copy-move manipulation. The algorithm presents the outcomes by either displaying them or preserving them, frequently including the creation of bounding boxes around the identified corresponding areas. Figure 4 Shows the visual outcome of the detection on some images.

Additional steps that can be included include threshold adjustment and assessment, where the matching threshold can be adjusted based on the properties of the dataset, and the success of the algorithm can be evaluated using metrics such as precision, recall, and F1. Moreover, the procedure can be iterated for many images when dealing with a dataset. To summarise, the technique utilizes a combination of multiscale key-point detection and descriptor learning to systematically and effectively detect instances of copy-move forgeries in digital images. The algorithm 1 provides the algorithmic stages for the proposed approach.

VI. EVALUATION METRIC AND DATASET

The tests done in our study were centered around the assessment of functionality at each pixel. The True Positive metric (TP) was employed to represent the overall count of pixels identified as manipulated that are indeed manipulated. The False Positive metric (FP) was utilized to represent the total count of pixels inaccurately identified as manipulated. Conversely, the False Negative metric (FN) was employed to represent the total count of pixels inaccurately identified as not manipulated. Utilizing the above values, we computed the metrics of True Positive Rate (TPR) or Recall (R),

FIGURE 4. Stages of manipulation detection. From left to right, the images include the original, manipulated, ground truth, detected keypoint, and detected copy-move regions.

TABLE 1. Overview of the dataset employed for experimental work.

Dataset	Total Image	Size of Image	Content of Image	Format of Image
CoMoFoD	2160	3000X2000 512X512	People, Urban Landscapes, Transportation Vegetation, Wall, Building.	JPEG PNG
CMFD	1632	388x2592 800x533	Outdoor spots, wildlife, Structure, Indoor Setting.	PNG
CASIA II.0	12323	800x600 320x240	Indoor Scenery, Plant Animal, Architecture.	BMF, TIFF JPEG
GRIP	240	1024X768	Birds, Floral, Animals Construction, Arid, People, Sky.	PNG
MICC-600	600	2048x1536	Flowra, Animals, People Home, Natural landscapes	JPEG
MICC-220	220	737x492	Road, People, Tree Shelves, Beach, Mountain	JPEG
COVERAGE	200	400X486	Indoor view, Rooms, Stores Public places, objects	TIFF

F1, and Precision (P) [12]. The F1 serves as the principal assessment statistic employed to evaluate the efficacy of our proposed methodologies and to make comparisons with other documented techniques [24]. The value of the variable in question is measured on a scale from 0 to 1, with 1 representing the highest quality and 0 representing the lowest quality. In our research study, we have converted this value into a percentage. The associations among TPR or R, F1, and P with TP, FP, and FN can be described as:

$$R* = TPR = \frac{TP}{TP + FN},$$
$$P* = \frac{TP}{TP + FP},$$
$$F1* = \frac{2TP}{2TP + (FP + FN)}.$$

Multiple evaluation criteria, including P, R or TPR, and F1, were employed to assess the effectiveness of our approach. Precision is a measure that calculates the proportion of accurately determined positive detections out of the total number of positive detections. The concept of Recall refers to the proportion of accurately detected positive detection to the total number of genuinely manipulated areas. The F1 is calculated as the average of the harmonics of Precision and Recall, providing a fair evaluation of the algorithm's total efficacy. The aforementioned measures offer significant insights into the Precision and efficacy of our approach in identifying instances of copy-move manipulation. These metrics provide a comprehensive assessment of our method's performance and are particularly useful for comparative analyses with other reported techniques in the field.

A. TYPE OF MANIPULATIONS CONSIDERED

The categorization of copy-move manipulation can be classified into four primary categories, which are determined by the specific technique employed in the creation of the manipulation. The subsequent items are:

- The act of duplicating and relocating part of an image within the same image can be classified as a simple copy-move manipulation. This could be one instance or many instances of a particular duplicated region. This is known as multiple copy-move manipulations. These created images have not undergone any post-processing.

Algorithm 1 Multiscale Key-Point Detection and Descriptor Learning

1: Input: Digital image suspected of containing copy-move forgeries.
2: Output: Detected key-points and potential matching regions.
3: Load the digital image.
4: Apply multiscale key-point detection network.
5: Extract patches centered on detected key-points.
6: Train descriptor network using triplet loss.
7: Use trained descriptor network to compute embeddings for detector key-point.
8: Calculate pairwise distances between embeddings for matching threshold.
9: Define matching threshold.
10: Find key-point pairs with distances below the threshold for identifying matching key-points.
11: Display or save results, draw bounding boxes for visualization results.
12: Fine-tune matching threshold, evaluate performance.
13: If working with a set of images, repeat the process.

- We did a study that included manipulated images that were post-processed to conceal their manipulations. In such cases, post-processing is frequently performed with a high level of expertise to make it more challenging to detect manipulations. As part of our study, we looked at what happened to these changed images when we changed the JPEG compression level and added noise.
- Aside from the basic post-processing method used to hide manipulation evidence, there are also auxiliary processes that help hide manipulation evidence even more. The detection of such techniques poses challenges due to their ability to consistently modify image pixels. The techniques encompassed in this set include Brightness adjustment (BC), Contrast adjustment (CA), Colour reduction (CR), and Blur.
- Geometric transformations have gained popularity as a technique for producing manipulated images due to their ability to make persuasive copy-move manipulation. These modifications can be implemented using three distinct methods: There are three primary methods for manipulating regions in a geometric context: 1) the duplication of a region followed by rotation, 2) the duplication of a region followed by scaling and translation, and 3) the combination of scaling and rotation on a duplicated region. Images that have been rotated 180 degrees are what we are most interested in.

B. DATASET

We have utilized seven publicly available datasets in our study, each serving as a valuable resource for evaluating our proposed method. These datasets include CMFD [33], GRIP [34], CoMoFoD [35], MICC-600 [36], MICC-220 [36], CASIA II.0 [37] and COVERAGE [38]. To provide a comprehensive overview of these datasets, we have compiled various dataset-related details, as presented in Table (1).

1) CoMoFoD: There are 200 base images in the CoMoFoD dataset. There are a total of 160 altered images, 40 each for translation, rotation, scaling, combination, and distortion. In addition, every single altered image goes through a series of post-editing procedures that include things like compression, noise, color reduction, brightness adjustment, contrast augmentation, and blur.
2) CASIA II.0: There are a total of 7491 images in the collection, 5123 of which have been manipulated in some way. We chose 3274 copy-move manipulated images that had been altered in various ways from this collection. Images are transformed using translation, rotation, and scale; some are further processed employing JPEG compression and blurring of edges.
3) CMFD: There are a total of 48 texture-rich base images included in the CMFD dataset. Copy-move manipulation is used in these images, which includes translation, rotation, scaling, and any combination thereof. Combinations of image rotation, scaling, and JPEG compression, as well as other post-processing processes like Additive White Gaussian Noise (AWGN), are used to simulate attacks. Using a 2*circ* rotation, a 1% scaling, and a JPEG compression level of 80, we were able to build up a combination attack. Later configurations (4*circ*, 3%, 75), (6*circ*, 5%, 70), and (8*circ*, 7, 65) enhanced rotation and scaling but decreasing JPEG quality.
4) MICC-600: Approximately 440 of the 600 images in the MICC-600 dataset have not been altered in any way. Copy-move manipulation is used to construct the altered images, which can use any combination of translation, rotation, and scale. They also use methods like JPEG compression and noise for post-processing.
5) MICC-220: There are a total of 220 images with ground truth in the MICC-220 dataset, consisting of 110 original images and their corresponding altered variants. In this data set, manipulation is accomplished through the use of the copy-move method in conjunction with either translation, rotation, scaling, or all three. JPEG compression and noise are also used as post-processing operations.
6) COVERAGE: The original, manipulated, and ground truth images of the same object are all included in the 200-image Coverage dataset. Images from both indoor and outdoor settings are included. There are six different types of post-processing in the dataset. They are translation, rotation, scaling, changing the lighting, free form, and any mix of these five. In addition, twenty photographs have been altered in various ways using copy-move techniques.

The images from different datasets that were used for training and tests in copy-move manipulation detection are

TABLE 2. Information of dataset and number of images used for training and test phase of the proposed work.

Dataset	CoMoFoD	CMFD	CASIA II.0	MICC 600	MICC 220	COVERAG	GRIP
Training Images	1100	850	850	550	110	100	120
Test Images	380	782	550	450	110	100	120

shown in Table (2). The number of images from each dataset that were randomly selected to train the copy-move manipulation detection model is listed in the ''Training Images'' row of the Table (2). In the same way, the ''Test Images'' row shows how many images were used to see how well the trained model worked with new, unknown data. The selection of training and test sets for each dataset is typically made based on the principle of random sampling to ensure representative coverage of the data distribution. To ensure reliable experimentation with the limited amount of data, cross-validation was performed 10 times randomly. The final result was obtained by averaging the results from each cross-validation run and evaluated accordingly. In our study, we adopted this approach, drawing from seven varied datasets and employing random sampling of images from each dataset for both training and testing. This methodology was key to ensuring a thorough and robust evaluation of our approach.

VII. FINDINGS AND INTERPRETATION

The Multiscale detector exhibits the capacity to extract resilient and consistent key-points from images, despite its initial features having changed. The detection of potential instances of copy-move manipulation can be accomplished by conducting a comparison of the extracted key-points obtained from various regions of an image. Moreover, the utilization of the Multiscale Detector and descriptor allows for the establishment of correspondences among the extracted key-points, enabling a precise assessment of the extent of geometric alterations applied to the duplicated region. The aforementioned data can thereafter be employed to facilitate the localization and identification of occurrences of counterfeiting.

Through the implementation of our conducted tests and subsequent assessment, we have effectively showcased the effectiveness of the proposed approach in accurately detecting instances of manipulation within a wide range of image categories. This encompasses images that went through various manipulations and subsequent modifications. The criteria for assessment utilized in our study have shown that our technique attains notable levels of Precision, Recall, and F1. The results of this study demonstrate that the used methodology exhibits effectiveness in detecting occurrences of copy-move manipulations, proving its potential as a helpful tool in the domain of forensic image analysis.

All experiments were conducted using the OpenCV-Python public library on a computer system equipped with an Intel Core i7 processor, 8 GB RAM, and a CPU clocked at 2.80GHz. The Multiscale Detector's computational cost for copy-move tampering detection is primarily influenced by recursive convolutions, the scale-dependent branching mechanism, and the convolutional and pooling layers. Among these components, the convolutional layers and recursive convolutions are the most computationally intensive.

A. DETECTION OF SIMPLE COPY-MOVE

The focus of our study pertains to the detection and analysis of instances of simple copy-move manipulation within the entirety of the datasets including various images like. The study conducted experiments to investigate two situations of simple copy-move manipulation, namely multiple and single copy-move scenarios. The approach we provided demonstrated strong performance in both scenarios. In Figure (5), a selection of identified manipulations for both multiple and single copy-move scenarios has been included. The proposed methodology ensures precision by utilizing a diverse range of images that exhibit distinct textures.

The results achieved for simple copy-move manipulation detection are presented in Table (3). The findings exhibit a consistent level of performance across various datasets. The approach we present demonstrates superior performance compared to the standard key-point-based approach [21], [39] across multiple datasets including GRIP, CMFD, CASIA II.0, CoMoFoD, COVERAGE, MICC-600, and MICC-220. However, the results in the GRIP dataset exhibit a modest decrease. The aforementioned phenomenon can be ascribed to the existence of highly smooth visual representations and depictions featuring elaborate patterns, such as sculptures adorned with engravings. Insufficient key-point quantity in smooth images hinders the precise detection and location of manipulations. On the other hand, in images that possess intricate patterns and exhibit self-repeating structures, the presence of numerous comparable key-points can result in inaccurate identification of manipulated areas. When evaluating the effectiveness of the suggested methodology on certain datasets, it is crucial to take into account these elements.

In a similar vein, numerous experiments were conducted to detect and localize instances of copy-move manipulations. The findings are presented in Figure (5). Upon comparing these results with the previous key-point-based approach, it is evident that the Multiscale Detector-based strategy demonstrates greater efficiency.

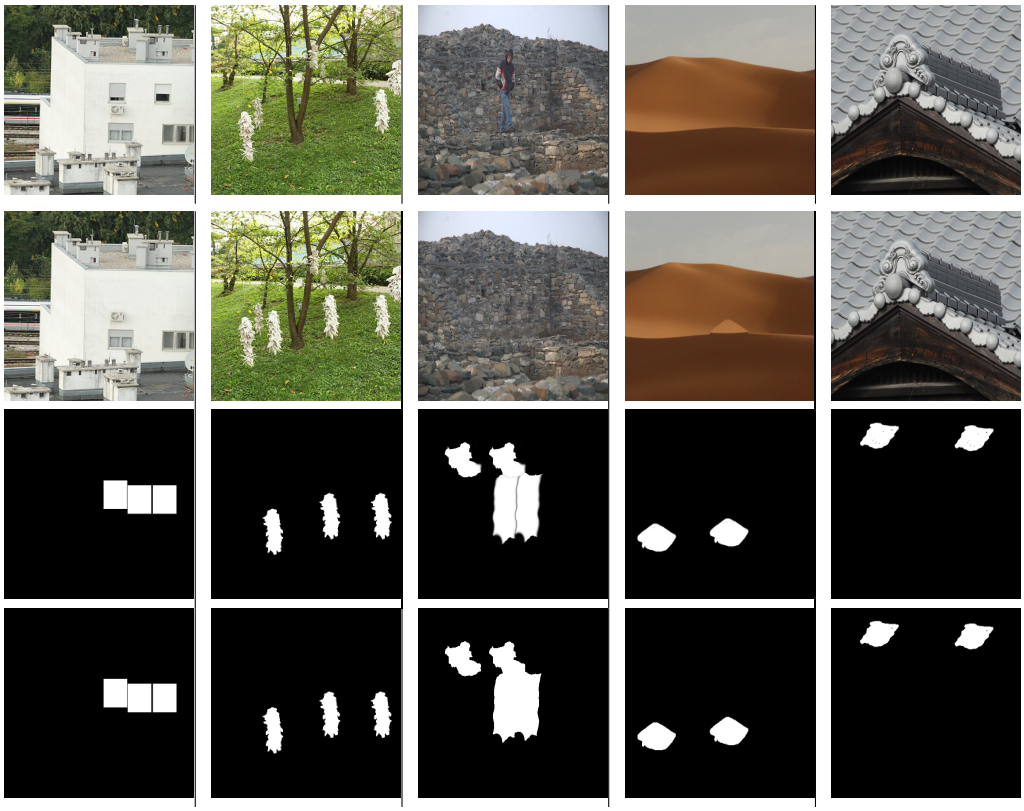

FIGURE 5. Results of manipulation detection for images. From top to bottom, the images include the original, manipulated, ground truth, and detected images.

TABLE 3. Comparative F1 analysis of our results with recent publications on simple copy-move manipulation in images.

Dataset	Zhu [26]	Bi [39]	Li [40]	Diwan [21]	Proposed
CMFD	–	92.87	98.91	97.61	**98.82**
CASIA II.0	45.52	–	–	95.36	**98.64**
CoMoFoD	–	–	–	98.43	**98.47**
COVERAGE	50.99	–	72.78	97.50	**98.03**
MICC-2000	–	–	91.50	97.14	**97.82**
MICC-600	–	–	91.50	97.14	**97.13**
GRIP	–	92.98	100	95.12	**97.08**
MICC-220	–	–	99.10	98.43	**96.87**

B. DETECTION OF MANIPULATED IMAGES WITH POST-PROCESSING

In instances where manipulation is executed with expertise, the modified image might need additional post-processing methods to conceal the act of manipulation. The present study analyzed various post-processing techniques, including but not limited to JPEG compression, Noise addition, color reduction, brightness change, contrast adjustment, and blur. The approach we have developed effectively identified and accurately determined a particular location of the manipulation, even in images that have undergone subsequent processing. These methods are frequently employed to conceal instances of copy-move manipulation in images.

The Multiscale Detector detector demonstrates the ability to extract stable and resilient key-points from post-processed images, even in cases where the original features have undergone modifications. Through the process of comparing the extracted key-points obtained from various regions within the image, the detector can identify possible instances of copy-move manipulation. Further, the utilization of the Multiscale Detector descriptor allows for the matching of the extracted key-points, enabling an accurate determination of the extent of geometrically transformed regions. The aforementioned data is effectively employed to correctly identify and pinpoint the location of the manipulation.

1) JPEG COMPRESSION

JPEG compression is a commonly employed technique for decreasing the size of the image while maintaining an acceptable level of visual quality. JPEG compression is a form of lossy compression employed to decrease the size of an image by selectively eliminating certain elements. The presence of these factors poses challenges in the identification of manipulated content within these images. The Multiscale Detector algorithm operates by identifying and characterizing the unique attributes present inside an image, including but not limited to edges, corners, and blobs. The identification of these features is accomplished by the utilization of a

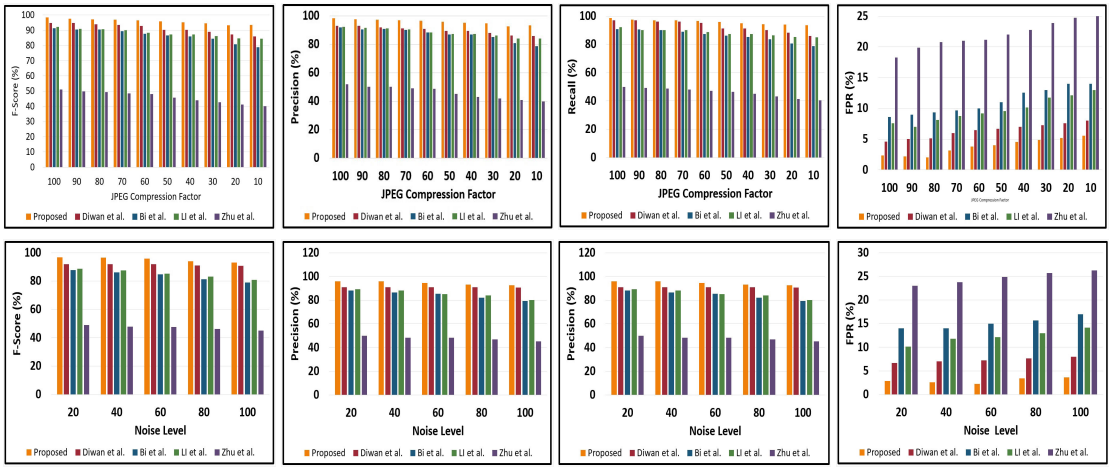

FIGURE 6. Comparative analysis of manipulation detection for images under varied levels of JPEG compression and additive noise, displaying precision, recall, F1, and True Positive Rate (TPR) from left to right with state-of-the-art approaches.

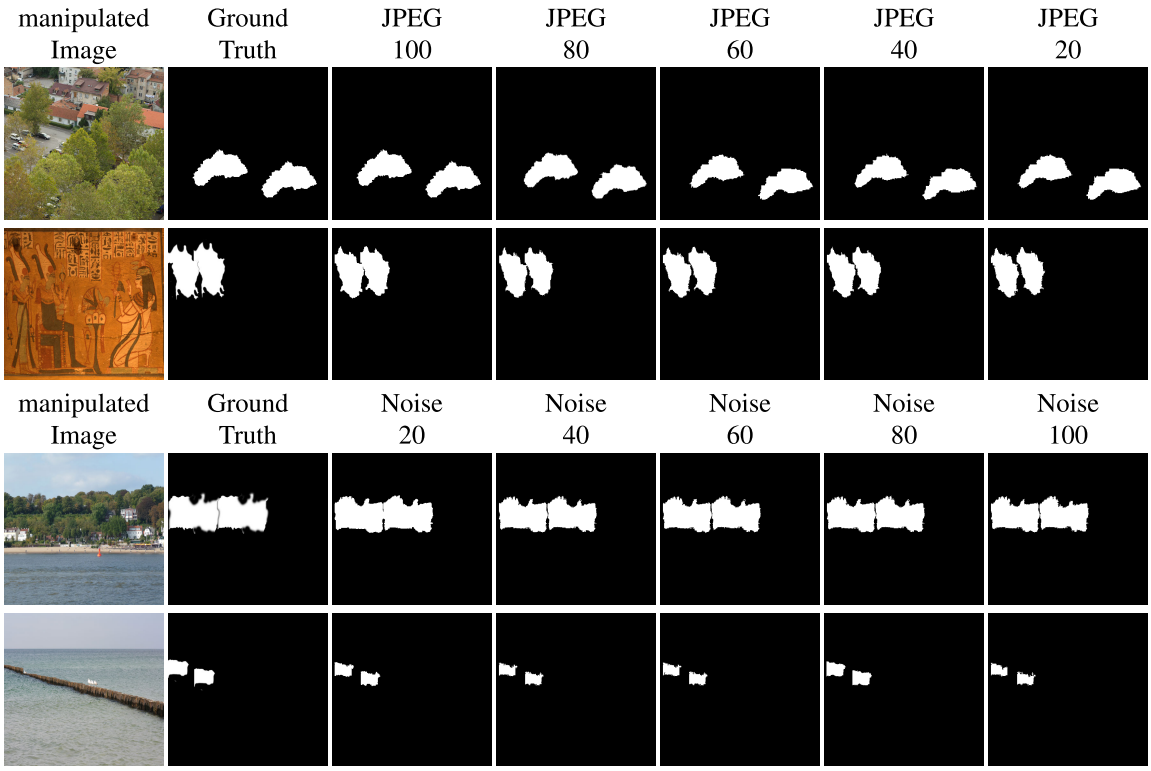

FIGURE 7. Manipulation detection result for images under different levels of JPEG compression and additive noise.

convolutional neural network, a computational model that is capable of learning and recognizing visual patterns within images across various scales.

To identify the presence of JPEG compression using the Multiscale Detector method, the initial step involves extracting key-points and descriptors from the image portions that have been copied and moved. This is achieved by utilizing the Multiscale Detector detector. Next, we proceed to identify the critical spots by employing a nearest-neighbor search.

In the conducted test, a series of digital images were employed, each exhibiting varying degrees of JPEG compression, spanning from JPEG100 (representing the minimal amount of compression) to JPEG20 (representing the maximal level of compression). The results exhibited a progressive decline as the compression level rose, as depicted in Table (4). As anticipated, the F1 exhibited a decline with more compression; still, commendable outcomes were attained across all instances. It is well acknowledged that as compression levels increase, the high-frequency components

TABLE 4. Comparative analysis of manipulation detection for images under different levels of JPEG compression and additive noise for two datasets CMFD and CoMoFoD, alongside CenSurE key-point based approach [21].

Attack Level	F1 Proposed CMFD	F1 Diwan [21] CMFD		F1 Proposed CoMoFoD	F1 Diwan [21] CoMoFoD
JPEG100	**98.52**	94.87		**98.38**	94.51
JPEG90	**97.53**	94.87		**97.75**	93.92
JPEG80	**97.06**	93.96		**96.99**	93.75
JPEG70	**96.94**	93.61		**96.63**	92.80
JPEG60	**96.53**	92.88		**95.99**	91.77
JPEG50	**95.89**	90.18		**95.49**	90.53
JPEG40	**95.13**	90.18		**94.17**	90.53
JPEG30	**94.51**	89.01		**93.89**	89.45
JPEG20	**93.37**	87.23		**92.39**	87.31
Noise20	**96.85**	91.90	Noise1	**96.51**	90.89
Noise40	**96.43**	91.90			
Noise60	**95.75**	91.90	Noise2	**95.49**	90.36
Noise80	**93.99**	91.01			
Noise100	**93.19**	90.88	Noise3	**93.22**	89.91

TABLE 5. Comparative analysis manipulation detection for images under different levels of angle rotation for CMFD dataset, alongside CenSurE key-point based approach [21].

Angle Rotation	F1 Proposed CMFD	F1 CenSurE CMFD	Angle Rotation	F1 Proposed CMFD	F1 CenSurE CMFD
2	**96.85**	94.91	20	**92.51**	89.88
4	**96.19**	94.05	40	**91.36**	87.65
6	**95.27**	94.99	60	**90.48**	87.65
8	**94.53**	93.90	180	**90.21**	84.88
10	**93.97**	93.88	–	–	–

of an image, including edges, corners, and gradients, undergo a steady process of smoothing. The key-point detector heavily depends on these high-frequency features as they play a critical role in detecting key-points that are both robust and stable.

The visual quality of compressed images employing JPEG compression is observed to deteriorate and exhibit blocking artifacts as the compression level falls below 40. The presence of these artifacts may have an impact on the identification of counterfeit elements within images that have undergone more extensive compression. The Multiscale Detector architecture demonstrates the ability to mitigate the adverse effects of increased compression levels through its capacity to identify and analyze the low level characteristics of an image. Nevertheless, the influence of compression artifacts becomes apparent when examining the F1 and accuracy of localizing the manipulated parts.

2) ADDITIVE NOISE

The introduction of noise into an image results in the emergence of diverse corners and edges. As the amount of noise increases, it can cause a blurring effect. The presence of various unrelated edges can potentially hinder the accurate recognition of key-points. Multiscale descriptors play a pivotal role in maintaining consistent and reliable performance, particularly when faced with elevated levels of noise in an image. Their adaptability across different scales allows them to effectively capture both fine and coarse details, ensuring a robust feature extraction process. This multiscale nature contributes to the stability of the output by enabling the descriptor to focus on relevant information and discriminate between signal and noise. The integration of local and global context information, coupled with multiresolution analysis techniques, further enhances the descriptor's ability to provide a comprehensive and stable representation, making it well-suited for scenarios where noise may impact traditional descriptors.

The outcomes for varying degrees of additive noise are presented in Table (4). The table clearly demonstrates a decline in the F1 as the level of noise in the images increases. The occurrence of this phenomenon can be attributed to the limited number of key-points present in the image, which subsequently impacts the accurate identification and localization of the manipulated region. However, despite the presence of increased amounts of noise, we managed to successfully identify the manipulated parts within the images.

The Multiscale Detector can identify and accurately locate instances of manipulation in compressed as well as noisy images by doing a comparative analysis of the unique characteristics present in the original and modified image regions. In Figure 6, a comparative graph illustrates the results of the proposed method for JPEG compression, alongside existing approaches such as Zhu et al. [26], Bi and Pun [39], Li and Zhou [40], and Diwan et al. [21]. The results of the proposed approach showcase superior performance when compared to recent published works. Key-point analysis plays a crucial role in detecting copy-move manipulation, particularly in the presence of post-processing attacks like JPEG compression and Additive noise. The graph visually emphasizes the effectiveness of the proposed method, highlighting its competitive edge over other state-of-the-art techniques.

3) ADDITIONAL PROCESSING

The concealment of copy-move manipulation can be achieved by the utilization of diverse post-processing techniques that manipulate the pixel-level characteristics of an image, effectively obscuring any discernible evidence of the manipulation. Various processes encompass contrast modification, color reduction, brightness alteration, and blurring. An example of this phenomenon is that when the brightness of the manipulated image is heightened, it diminishes the contrast value and thus increases the occurrence of false negatives. As a result, the recall rate and overall accuracy of detection (F1) are reduced. In contrast, the process of color reduction entails decreasing the intensity level across all color channels, causing multiple hues to be represented by identical values. This, in turn, amplifies the prominence of edges and has a consequential impact on the accuracy of detection.

TABLE 6. Comparative analysis of manipulation detection for images under different levels of scale change for CMFD dataset, alongside CenSurE key-point based approach [21].

Sacle Factor	F1 Proposed	F1 CenSurE
2	**96.88**	93.49
4	**95.65**	93.95
6	**95.38**	92.46
8	**93.89**	92.01
10	**93.73**	92.88

The efficacy of our suggested approach, which is based on a Multiscale Detector, lies in its ability to accurately identify and pinpoint instances of manipulation within images that have undergone diverse post-processing procedures. Experiments were done on several post-processing techniques using the CoMoFoD dataset. The data indicate that our approach consistently detects and precisely determines the location of manipulation, regardless of the degree of post-processing used, ranging from mild (level 1) to severe (level 3). The outcomes of additional post-processing are depicted in Figure (8). The Multiscale Detector detector employs a clustering technique to preserve similarity across key-points, while also effectively capturing local image information using a Multiscale Detector.

C. DETECTION OF MANIPULATED IMAGES WITH GEOMETRICAL TRANSFORMATIONS

The task of identifying image manipulation becomes increasingly complex when the replicated portion undergoes geometric alterations before its relocation. In instances of this nature, the relationship between the copy and move parts undergoes substantial modification, particularly when the extent of rotation and scale is considerable. The task of extracting similar features from the parts that have gone through substantial geometric modifications poses greater challenges compared to circumstances when the transformations are minimal.

The task of identifying instances of copy-move manipulations in images that have gone through geometric transformations poses a significant challenge because of the lack of invariance of some image features to rotation and scaling. To tackle this issue, we employed the Multiscale Detector key-point detector, a reliable method for obtaining consistent image key-points that exhibit resilience against rotation and scaling. The utilization of homography adaption in the Multiscale Detector algorithm proves to be advantageous within the domain of copy-move manipulation detection inside geometrically altered images. This technique facilitates the identification and alignment of key-points, enabling the identification and matching of these key-points, especially in cases where the manipulated region has experienced various geometrical transformations, such as rotation, scaling, or a combination of transformations.

When a manipulated part of an image experiences a homographic transformation, the corresponding key-points in the original and manipulated regions will not exhibit direct correspondence. Nevertheless, through the process of calculating the homography matrix across the two parts of the image, it becomes possible to transfer the key-points from the original part to the corresponding coordinates in the manipulated part of the image. Consequently, these key-points can then be matched with the key-points that were discovered in the manipulated region. The proposed method enables the identification of copy-move manipulation in images that have been subjected to homographic transformations.

The process of homography adaption in a Multiscale Detector entails the estimation of the homography matrix between the key-points found in both the original and manipulated parts. This estimation is achieved through the use of the RANSAC algorithm. The homography matrix can subsequently be employed to convert the key-points inside the initial region into the corresponding coordinates within the manipulated part, hence facilitating the establishment of correspondence between the two regions.

The repeatability of the Multiscale Detector is a beneficial characteristic when it comes to the identification of image manipulation including rotation. The geometrical invariance, in conjunction with transformation computation, enables the detection of manipulations even in cases where the copy-move part experiences a change in scale, either from tiny to big or from big to tiny. In the proposed approach, the advantages of the Multiscale Detector was exploited by utilizing its feature extraction capabilities. This enables us to preserve correspondence across copy-move images, even when they undergo rotations of varying magnitudes.

The proposed methodology has demonstrated a notable enhancement in the detection precision for rotation in comparison to the key-point-based approach previously discussed by Diwan et al. [21]. The method demonstrates notable efficacy in identifying instances of manipulation that include substantial angle rotations, specifically those of 20°, 40°, 60°, and 180°. The findings of the proposed study exhibit steady performance across several small angles of rotation, specifically 2°, 4°, 6°, 8°, and 10°. Figure (9) presents a collection of images exhibiting varying rotation angles.

The Table (6), displays the outcomes of detecting and localizing copy-move manipulated images that have undergone scale changes. Various datasets contain manipulated images that incorporate both rotation and scaling techniques. Our approach effectively identifies and localizes instances of manipulations in various types of attacks. The combination of rotation and scaling is exemplified by the images presented in Figure (10).

D. LIMITATIONS AND FUTURE SCOPE

Processing multiscale information may be more computationally complex than single-scale methods, which could be problematic in situations when resources are restricted or in real-time applications. Additionally, small-scale changes

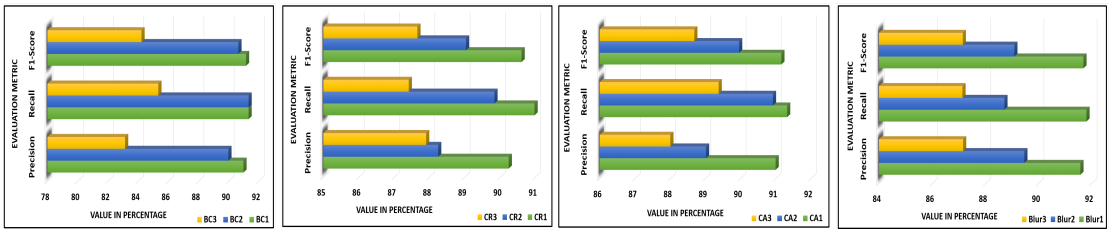

FIGURE 8. Copy-move manipulation detection results for images with additional post-processing, e.g., Brightness change (BC), Colour reduction (CR), Contrast adjustment (CA), and blur for the CoMoFoD dataset.

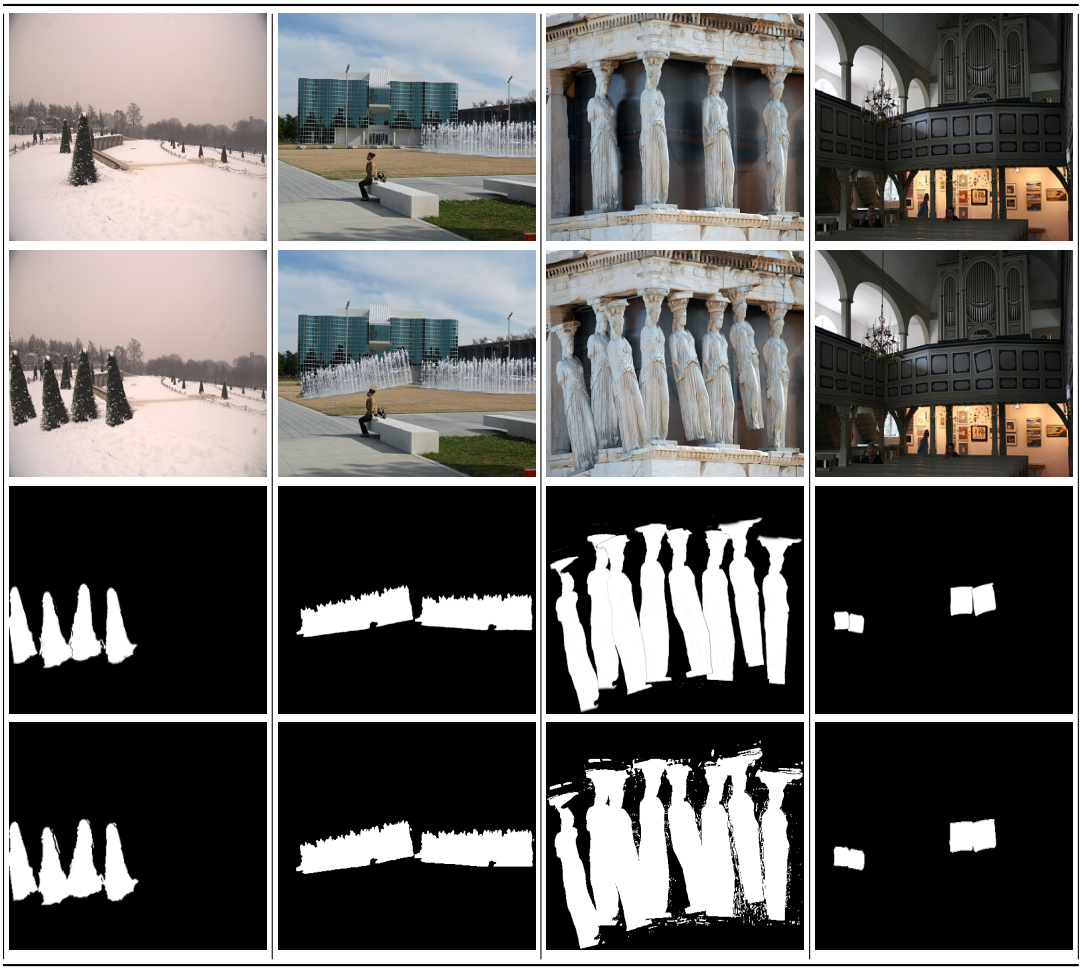

FIGURE 9. Results of manipulation detection for images subjected to angle rotation from left to right, the images include the original, manipulated, ground truth, and detected images.

may be difficult for multiscale detectors to precisely detect, especially in instances where the copied region is heavily transformed.

Multiscale copy-move detection methods are expected to evolve in the future to better handle emerging challenges in digital image forensics other than copy-move tampering. One approach to enhance feature representation and detection accuracy across various dimensions is to utilize advanced deep learning architectures like graph neural networks or attention mechanisms. Additionally, research may focus on creating adaptive multiscale techniques that can automatically adjust scale selection based on image characteristics and types of manipulation. Another approach to detecting sophisticated copy-move tampering could involve exploring new methods for handling complex transformations, such as non-rigid deformations or content-aware manipulations. Additionally, combining multiscale detectors with forensic techniques like image hashing or content-based retrieval could lead to more comprehensive and efficient forgery detection systems.

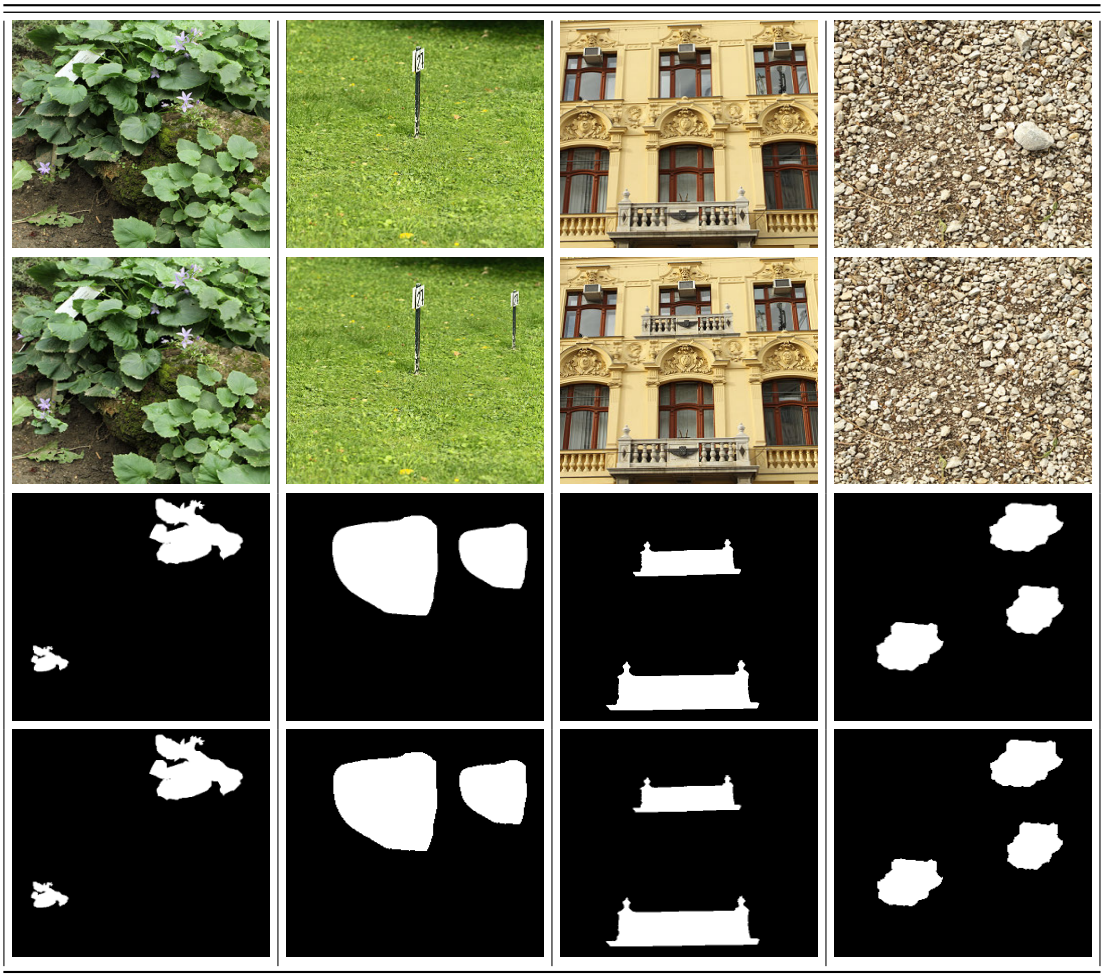

FIGURE 10. Results of copy-move tampering detection for images subjected to scale change and combined attacks, from left to right, the images include the original, manipulated, ground truth, and detected images.

VIII. CONCLUSION

The Multiscale Detector architecture uses scale-space analysis, key-point localization, subpixel refinement, and neural network-based descriptor learning. This combination detects robust and unique key-points and extracts proper feature descriptions. This approach helps the system acquire and interpret local image structures and patterns, making it ideal for copy-move modification detection. The proposed copy-move fusion method, which is capable of being trained from start to finish, utilizes the Multiscale Detector key-point detector and descriptor's positive attributes. This approach finds and pinpoints digital imagine copy-move manipulation. Data-driven algorithms update their learning process using training data to improve manipulation detection.

The Multiscale Detector feature detection method demonstrates exceptional performance in a variety of copy-move manipulation scenarios, including simple copy-move instances (both single and multiple), post-processed copy-move cases involving JPEG compression and noise, geometrically transformed copy-move instances involving angle rotation and scale modification, and additional image manipulations such as brightness modification, color reduction, contrast adjustment, and blur. The inherent robustness of our methodology enables the efficient identification of manipulated instances among a diverse array of copy-move-manipulated images. In addition, our research emphasizes the incorporation of images with diverse textures, encompassing both smooth and coarse images with dense textures. This deliberate strategy improves the adaptability of our detection method, allowing it to detect a wide variety of manipulated images.

Our proposed solution has been rigorously evaluated on seven open-source datasets, and the results confirm its superior performance compared to prior image manipulation detection methods. In addition, the algorithm's capacity for efficient data analysis makes it well-suited for applications requiring rapid detection of manipulations. In general, our methodology produces consistent and reliable results when applied to various categories of image manipulations, highlighting its efficacy and adaptability in identifying copy-move manipulations.

ACKNOWLEDGMENT

Rajesh Mahadeva thanks Marwadi University for allowing them to use their facility.

REFERENCES

[1] M. Muchmore. (2018). *The Best Photo Editing Software of 2018*. [Online]. Available: https://www.pcmag.com/article2/0

[2] BBC. (2021). *How Fake Images Change Our Memory and Behaviour*. [Online]. Available: https://www.bbc.com/future/article/20121213-fake-pictures-make-real-memories

[3] P. Maji, M. Pal, R. Ray, and R. Shil, "Image tampering issues in social media with proper detection," in *Proc. 8th Int. Conf. Rel., INFOCOM Technol. Optim.*, Jun. 2020, pp. 1272–1275.

[4] J. Andrews. (2021). *The Hidden Fingerprint Inside Your Photos*. [Online]. Available: https://www.bbc.com/future/article/20210324-the-hidden-fingerprint-inside-your-photos

[5] H. Farid, *Fake Photos*. Cambridge, MA, USA: MIT Press, 2019.

[6] G. Tahaoglu, B. Ustubioglu, G. Ulutas, M. Ulutas, and V. V. Nabiyev, "Robust copy-move forgery detection technique against image degradation and geometric distortion attacks," *Wireless Pers. Commun.*, vol. 131, no. 4, pp. 2919–2947, Aug. 2023.

[7] A. Diwan and U. Sonkar, "Visualizing the truth: A survey of multimedia forensic analysis," *Multimedia Tools Appl.*, pp. 1–28, Oct. 2023. [Online]. Available: https://link.springer.com/article/10.1007/s11042-023-17475-3

[8] B. Ahmed, T. A. Gulliver, and S. alZahir, "Blind copy-move forgery detection using SVD and KS test," *Social Netw. Appl. Sci.*, vol. 2, no. 8, pp. 1–12, Aug. 2020.

[9] G. Gopakumar, "A survey on image splice forgery detection and localization techniques," in *Proc. 3rd Int. Conf. Intell. Comput. Instrum. Control Technol. (ICICICT)*, Aug. 2022, pp. 1242–1247.

[10] X.-Y. Wang, C. Wang, L. Wang, H.-Y. Yang, and P.-P. Niu, "Robust and effective multiple copy-move forgeries detection and localization," *Pattern Anal. Appl.*, vol. 24, no. 3, pp. 1025–1046, Aug. 2021.

[11] S. B. G. T. Babu and C. S. Rao, "Copy-move forgery verification in images using local feature extractors and optimized classifiers," *Big Data Mining Analytics*, vol. 6, no. 3, pp. 347–360, Sep. 2023.

[12] A. Diwan, V. Mall, A. Roy, and S. Mitra, "Detection and localization of copy-move tampering using features of locality preserving projection," in *Proc. 5th Int. Conf. Image Inf. Process. (ICIIP)*, Nov. 2019, pp. 397–402.

[13] K. M. Hosny, H. M. Hamza, and N. A. Lashin, "Copy-for-duplication forgery detection in colour images using QPCETMs and sub-image approach," *IET Image Process.*, vol. 13, no. 9, pp. 1437–1446, Jul. 2019.

[14] G. Gani and F. Qadir, "A robust copy-move forgery detection technique based on discrete cosine transform and cellular automata," *J. Inf. Secur. Appl.*, vol. 54, Oct. 2020, Art. no. 102510.

[15] X.-Y. Wang, C. Wang, L. Wang, L.-X. Jiao, H.-Y. Yang, and P.-P. Niu, "A fast and high accurate image copy-move forgery detection approach," *Multidimensional Syst. Signal Process.*, vol. 31, no. 3, pp. 857–883, Jul. 2020.

[16] A. Diwan, D. Kumar, R. Mahadeva, H. C. S. Perera, and J. Alawatugoda, "Unveiling copy-move forgeries: Enhancing detection with superpoint keypoint architecture," *IEEE Access*, vol. 11, pp. 86132–86148, 2023.

[17] N. Kumar and T. Meenpal, "Salient keypoint-based copy–move image forgery detection," *Austral. J. Forensic Sci.*, vol. 55, no. 3, pp. 331–354, May 2023.

[18] S. I. Lee, J. Y. Park, and I. K. Eom, "CNN-based copy-move forgery detection using rotation-invariant wavelet feature," *IEEE Access*, vol. 10, pp. 106217–106229, 2022.

[19] A. Kalluvilayil Venugopalan and G. Gopakumar, "Keypoint-based detection and region growing-based localization of copy-move forgery in digital images," in *Computer Vision and Machine Intelligence*. Cham, Switzerland: Springer, 2023, pp. 513–524.

[20] X.-Y. Wang, X.-Q. Wang, P.-P. Niu, and H.-Y. Yang, "Accurate and robust image copy-move forgery detection using adaptive keypoints and FQGPCET-GLCM feature," *Multimedia Tools Appl.*, vol. 83, no. 1, pp. 2203–2235, Jan. 2024.

[21] A. Diwan, R. Sharma, A. K. Roy, and S. K. Mitra, "Keypoint based comprehensive copy-move forgery detection," *IET Image Process.*, vol. 15, no. 6, pp. 1298–1309, 2021.

[22] Z. Su, M. Li, G. Zhang, Q. Wu, M. Li, W. Zhang, and X. Yao, "Robust audio copy-move forgery detection using constant Q spectral sketches and GA-SVM," *IEEE Trans. Dependable Secur. Comput.*, vol. 20, no. 5, pp. 1–15, Oct. 2022.

[23] C. Wang, Z. Zhang, Q. Li, and X. Zhou, "An image copy-move forgery detection method based on SURF and PCET," *IEEE Access*, vol. 7, pp. 170032–170047, 2019.

[24] Y. Zhang, G. Zhu, X. Wang, X. Luo, Y. Zhou, H. Zhang, and L. Wu, "CNN-transformer based generative adversarial network for copy-move source/target distinguishment," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 33, no. 5, pp. 2019–2032, May 2023.

[25] C. Wang, Z. Huang, S. Qi, Y. Yu, G. Shen, and Y. Zhang, "Shrinking the semantic gap: Spatial pooling of local moment invariants for copy-move forgery detection," *IEEE Trans. Inf. Forensics Security*, vol. 18, pp. 1064–1079, 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10007894

[26] Y. Zhu, C. Chen, G. Yan, Y. Guo, and Y. Dong, "AR-net: Adaptive attention and residual refinement network for copy-move forgery detection," *IEEE Trans. Ind. Informat.*, vol. 16, no. 10, pp. 6714–6723, Oct. 2020.

[27] S. B. G. Tilak Babu and C. Srinivasa Rao, "Statistical features based optimized technique for copy move forgery detection," in *Proc. 11th Int. Conf. Comput., Commun. Netw. Technol. (ICCCNT)*, Jul. 2020, pp. 1–6.

[28] K. H. Rhee, "Generation of novelty ground truth image using image classification and semantic segmentation for copy-move forgery detection," *IEEE Access*, vol. 10, pp. 2783–2796, 2022.

[29] M. Maashi, H. Alamro, H. Mohsen, N. Negm, G. Pasha Mohammed, N. Abdelaziz Ahmed, S. Saadeldeen Ibrahim, and M. Ibrahim Alsaid, "Modeling of reptile search algorithm with deep learning approach for copy move image forgery detection," *IEEE Access*, vol. 11, pp. 87297–87304, 2023.

[30] A. H. Khalil, A. Z. Ghalwash, H. Abdel-Galil Elsayed, G. I. Salama, and H. A. Ghalwash, "Enhancing digital image forgery detection using transfer learning," *IEEE Access*, vol. 11, pp. 91583–91594, 2023.

[31] J. Rao, S. Teerakanok, and T. Uehara, "ResTran: Long distance relationship on image forgery detection," *IEEE Access*, vol. 11, pp. 120492–120501, 2023.

[32] H. Altwaijry, A. Veit, and S. Belongie, "Learning to detect and match keypoints with deep architectures," in *Proc. Brit. Mach. Vis. Conf.*, 2016, pp. 1–14.

[33] V. Christlein, C. Riess, J. Jordan, C. Riess, and E. Angelopoulou, "An evaluation of popular copy-move forgery detection approaches," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 6, pp. 1841–1854, Dec. 2012.

[34] D. Cozzolino, G. Poggi, and L. Verdoliva, "Copy-move forgery detection based on PatchMatch," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2014, pp. 5312–5316.

[35] D. Tralic, I. Zupancic, S. Grgic, and M. Grgic, "CoMoFoD—New database for copy-move forgery detection," in *Proc. ELMAR*, Sep. 2013, pp. 49–54.

[36] I. Amerini, L. Ballan, R. Caldelli, A. Del Bimbo, and G. Serra, "A SIFT-based forensic method for copy–move attack detection and transformation recovery," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 1099–1110, Sep. 2011.

[37] J. Dong, W. Wang, and T. Tan, "CASIA image tampering detection evaluation database," in *Proc. IEEE China Summit Int. Conf. Signal Inf. Process.*, China, Jul. 2013, pp. 422–426.

[38] B. Wen, Y. Zhu, R. Subramanian, T.-T. Ng, X. Shen, and S. Winkler, "Coverage—A novel database for copy-move forgery detection," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2016, pp. 161–165.

[39] X. Bi and C.-M. Pun, "Fast copy-move forgery detection using local bidirectional coherency error refinement," *Pattern Recognit.*, vol. 81, pp. 161–175, Sep. 2018.

[40] Y. Li and J. Zhou, "Fast and effective image copy-move forgery detection via hierarchical feature point matching," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 5, pp. 1307–1322, May 2019.

ANJALI DIWAN (Senior Member, IEEE) received the Ph.D. degree from the Dhirubhai Ambani Institute of Information and Communication Technology (DAIICT), Gandhinagar, Gujarat. She is currently a Faculty Member with the CE-AI/Big Data Department, Marwadi University, Rajkot, Gujarat, India. She is a highly experienced academic and software industry professional with over 19 years of expertise. Her areas of academic and research interests include machine learning, image processing, artificial intelligence, deep learning, data security, multimedia forensics, and the application of technologies to address humanitarian challenges. She also serves as a member for the SAC Team of IEEE R10 (2023–2024) and the Section Chair for the IEEE Young Professionals Affinity Group of Gujarat Section (2022–2024).

RAJESH MAHADEVA (Member, IEEE) received the B.E. degree in electronics and instrumentation engineering from the Samrat Ashok Technological Institute (SATI), Vidisha, Madhya Pradesh, India, in 2006, the M.Tech. degree in control and instrumentation engineering from the Dr. B R Ambedkar National Institute of Technology (NIT), Jalandhar, Punjab, India, in 2009, and the Ph.D. degree from the Department of Polymer and Process Engineering, Indian Institute of Technology (IIT) Roorkee, Uttarakhand, India, in 2022. From 2011 to 2017, he was an Assistant Professor with the Technocrats Institute of Technology (TIT), Bhopal, and Marwadi University (MU), Rajkot, India. He is currently a Research Scientist with the Department of Physics, Khalifa University of Science and Technology, Abu Dhabi, United Arab Emirates. His research interests include modeling, simulation, optimization, and control of desalination and water treatment plants/processes using artificial intelligence techniques.

VINAY GUPTA (Member, IEEE) is currently an Assistant Professor with the Department of Physics, Khalifa University, Abu Dhabi, United Arab Emirates. His research interests include solid state physics, organic and perovskite solar cells, supercapacitors, lithium-ion batteries, and water filtration. He was a recipient of several international and national awards, including the Prestigious Shanti Swarup Bhatnagar (SSB) Award (2017) and the Thomson Reuters India Citation Award (2015).