International Journal of Engineering Trends and Technology
ISSN: 2231 – 5381 / https://doi.org/10.14445/22315381/IJETT-V70I10P225

Volume 70 Issue 10, 262-270, October 2022
© 2022 Seventh Sense Research Group®

Original Article

Deep Studying Signature for Obstruction obscure in Copy Move Image Forgeries

S. Shashikala[1], G. K. Ravikumar[2]

[1]*Department Computer Science, New Horizon College, Bangalore, Karnataka, India.*
[2]*Department of Computer Science and Engineering, BGS Institute of Technology, BG Nagara Mandya, Karnataka, India.*

[1]*Corresponding Author : shashi127@yahoo.com*

Received: 17 June 2022 Revised: 07 September 2022 Accepted: 12 October 2022 Published: 20 October 2022

Abstract - *Digital image tampering and counterfeiting can be done precisely with advanced photo editing tools available with various malicious intentions. It becomes necessary to verify the integrity of the image as images are becoming the information source in various computer-aided applications. Copy move counterfeits are created by copying a slice from one region of the image to another region. The current techniques for detecting copy-move counterfeits fail in the presence of partial occlusion or partial distortion created to falsify detection. This work proposes a deep learning signature to solve this problem. Deep learning signature is created using a probabilistic distribution function of occlusions. The coarse forgery regions are detected with scale-invariant feature transform-based keypoint matching. Deep learning signature matches the coarse forgery regions to detect the partially occluded copy move counterfeits.*

Keywords – *Copy-move forgery, Coarse forgery regions, Deep learning signature, Partially occluded counterfeit.*

1. Introduction

A digital image has become a most important information source in various fields like disease diagnosis, media, criminal forensics etc. Due to higher visual impact than text and creating a language-independent truthfulness about the event, images are preferred for communication. With the rapid use of images in various domains, there is also an increase in image counterfeiting for various malicious purposes like falsifying medical diagnoses, defaming people, disrupting social harmony, diverting criminal forensics etc. Image counterfeiting can be done to convey false impressions and create disastrous consequences [1-3]. With the availability of sophisticated tools, it becomes easy to create high-quality tampering that appears natural and authentic. Copy-move and splicing are the two most popular image tampering mechanisms. In case of copy-move forgery, a copy of an object is replicated to some other region in the image. In the case of splicing, an object in the image is pruned and replaced with another object.

The current methods proposed for detecting copy-move and slicing image forgery [4-25] work in three modes: frequency, spatial and hybrid. Spatial domain-based techniques use the information of statistical features of pixel and their locations to detect counterfeit regions. Frequency domain techniques use frequency analysis and wavelet transform feature analysis to detect counterfeit regions. The spatial and frequency domain techniques are used in different combinations of hybrid techniques. An important challenge in the current copy-move forgery detection technique is that it can be deceived easily by partial occlusion in the forgery regions. These partial occlusions make accurate localization of copy-move a challenge. This work considers this problem and proposes a deep learning signature-based copy move forgery detection. Deep learning signature is constructed for the objects as the probability distribution of objects in the presence of occlusions in the frequency domain. . The copy-move regions in the image are detected by combining the results of deformable coarse forgery region selection and deep learning signature tracking. Following are the contributions of this work.

(i) A novel deep learning signature with the probability of distribution of occlusion in frequency domain applying Quaternion Discrete Cosine Transform (QDCT) for detecting partially occluded copy move regions

(ii) Integration of deep learning signature-based detection with deformable model-based coarse forgery region selection for better speedup of copy forgery detection.

The Paper is organized as follows. The survey of existing techniques for image forgery detection is presented in Section II. The proposed deep learning model for fake detection in the presence of occlusions is presented in Section III. The results of the proposed work and its comparison to recent works are presented in Section IV. Finally, the concluding remarks and future work scope are given in Section V.

2. Survey

A hybrid technique combining discrete cosine transform (DCT) with local binary pattern (LBP) was proposed by Islam et al. [5]. DCT features are extracted from non-overlapping image blocks. LBP is applied to

DCT magnitude, and mean values are extracted as features. This hybrid feature is then matched over blocks to detect copied blocks. Combining DCT with LBP increased the fake detection accuracy to 95%, but the method is not transformation invariant. A hybrid technique specific to face copy move forgery was proposed by Cristin et al. [6]. This hybrid technique combined Gabor filer, wavelet and texture operator to extract features from the face and match it using an SVM classifier. But the method is very sensitive to illumination artifacts introduced into face copies. Guo et al. [7] used histograms for fake image detection. The approach is based on statistical differences between real and fake images in hue and saturation channels. But the method can be deceived easily by manipulating the hue and saturation values in the image. Li et al. [8] proposed a tampering region localization algorithm based on statistical features. Multiple detectors are used, and each detector's results are fused to get a tampering possibility map. All possible tampering regions are localized in the tampering possibility map. Higher false positives and the inability to localize in the presence of occlusion are the issues in this approach. Deep learning LSTM classifier, along with re-sampling features, was used to detect splicing forgeries by Bappy et al. [9]. LSTM used both spatial maps and frequency domain correlations. Though the method can provide pixel-wise predictions, it fails in the presence of occlusions. Chen et al. used Fractional Zernike moments features to detect copy-move forging [10]. The image is split into circular windows, and Zernike features are extracted from it. Windows are matched to identify similar patches. Zernike moments have a huge variance when occlusion is present, due to which certain copy regions are missed. Also, this method does not work for rotational transformations. Y.Li et al. [11] proposed a feature point-matching algorithm to detect copy-move forgery. Keypoints extracted from regions are matched iteratively to localize the tampered regions. The approach fails for even a small distortion in the copied regions. Mayer et al. [12] used the inconsistencies in lateral chromatic aberration (LCA) in the image to detect copy-move forgery. The image is split into grids. LCA estimates at the local and global levels are analyzed statistically to detect the forging. The approach works only for certain backgrounds and does not work for dynamic backgrounds. Offset guided searching is followed to detect copy-move forgery by Bi et al. [13]. Features extracted from different regions are matched to find any offset relation between features to detect copy-move regions. The offset method can be deceived easily with minor transformations. A key point-based copy move forgery detection was proposed by Wang et al. [14]. The image is first segmented using a superpixel segmentation algorithm. Keypoints are extracted and matched to detect copied regions. Though the method is resilient against transformation, it fails to match in the presence of occlusions. Similarly, Teerakanok et al. [15] used SURF and GLCM features to detect forgery. In the presence of occlusion, GLCM feature matching fails. Regions in the image are matched using the Gaussian operator to detect matches by Emam et al. [16]. Histogram features are extracted around the covariant key points in the image and matched to detect forged regions. But occlusions distort the localization of key points. Singh et al. [17] proposed a multi-modal approach for fake image detection. Correlation between the concepts expressed in the image to the metadata text is made to verify consistency; when consistency fails, the image is detected as copy. But this method fails in image copying. Zhang et al. [18] used deep learning and error-level analysis to detect fake images. Image compression ratios are measured in different image regions, and inconsistencies are detected as fake images. Ghoneim et al. [19] proposed a fake detection method for medical images. The method applies a noise map at different resolutions and looks for inconsistencies in the edge. But the method works only for edge inconsistencies during copying. Quaternion polar complex exponential transform (QPCET) for copy-move forgery detection by Thajeel et al. [20]. Invariant features are extracted from the image and matched using the KD-tree matching algorithm to detect copied regions. But detecting invariant features becomes difficult in the presence of occlusion. A convolutional kernel network (CKN) was used to detect copy-move forgeries by Liu et al. [21]. Image is segmented into regions, and regions are matched using the Convolutional kernel network. Though the method performs better than hand-crafted features, it cannot identify similarities in partially occluded images. Also, the method is not transformation invariant. Chou et al. [22] proposed Gabor filter-based copy-move forgery. The image is split into blocks, and block matching is done using Gabor filter features. The method can work only if objects fit in the block and are very sensitive to even a small distortion in shape. Mahmood et al. [23] used Wavelet features for the copy move forgery detection. Wavelet transform features were extracted from non-overlapping blocks of the image. Dimension reduction is done on the features. Features are then matched to detect copied blocks. But the method is not transformation invariant. Hosny et al. [24] detected copy-move forgery using exponential transform coefficients. The objects are segmented, and features are extracted from objects. These features are matched using Euclidean distance matching to detect copies. But the method shows a large Euclidean distance for occluded copies. Segnet Deep learning model trained with Haar wavelet features was used for copy move forgery detection by Khayeat et al. [25]. The objects in the images are segmented, and Haar wavelet features of level one decomposition are extracted from the objects. Objects match by passing the Haar wavelet features to the Segnet deep learning model. The method is not transformation invariant and fails in the presence of occlusions.

3. Deep Learning Signature-Based Fake Detection

The proposed solution is based on the assumption that fake colorized images have lower saturated colors and provides a preference for some color over others. It is difficult to observe this change through visual inspection. Analysis in the RGB domain is not effective in spotting these differences.

3.1. Coarse Region Selection

A deformable model [28] is used for region selection. In this model, curves and surfaces placed on the image are deformed by internal and external forces to fit the objects of interest. Among the different types of deformable models, Topological active net (TAN) [31] deforms mesh with the goal of energy minimization and fitting to the objects. Deformation is done by removing the links of mesh not fitting the object. Thresholding on the energy of the link is done to remove the link. The energy of the link is calculated as

$$\mathrm{E_{link}} = (\textstyle\sum_{\mathrm{p\varepsilon A}} \mathrm{DG_{evfc}(p)}.\frac{\mathrm{I(p)}}{\mathrm{I_{max}}}) \ / \ |\mathrm{A}| \qquad (1)$$

Where p is the pixel in the image of area A. I and $\mathrm{I_{max}}$ are the original image and the maximum intensity value. The mean energy of links is taken as the threshold for shunting down the links. Identification of holes starts from misplaced internal nodes. Setting the mean energy of the link as the threshold, the holes are created by cutting the links starting from misplaced internal nodes.

The energy ratio (r(n)) for an internal node is calculated in terms of external energy(Eext) and internal energy(Eint) as

$$\mathrm{r(n)} = \frac{\mathrm{E_{ext}(n)}}{\mathrm{E_{ext}(n)+\ E_{int}(n)}} \qquad (2)$$

The features of the scene are best represented in the external energy term.

$$\mathrm{E_{ext}}\big(\mathrm{v(m,n)}\big) = \ \omega \mathrm{f}\big[\mathrm{I}\big(\mathrm{v(m,n)}\big)\big] + \frac{\rho}{|\mathrm{N(m,n)}|}\sum\nolimits_{\mathrm{p\varepsilon N(m,n)}} \frac{1}{||\mathrm{v(m,n)-v(p)}||}\mathrm{f}\big[\mathrm{I}\big(\mathrm{v(p)}\big)\big] \qquad (3)$$

Where ω,ρ are the weights, $\mathrm{I}\big(\mathrm{v(m,n)}\big)$ is the value of the intensity of the pixel at position $\mathrm{v(m,n)}$. $\mathrm{N(m,n)}$ is the neighborhood of node at $\mathrm{(m,n)}$. Function f is defined as

$$\mathrm{f}\big[\mathrm{I}\big(\mathrm{v(m,n)}\big)\big] = \begin{cases} \mathrm{Imax} - \frac{{}^{\gamma}\,\overrightarrow{\mathrm{I}(\mathrm{v(m,n)})}}{\overrightarrow{\mathrm{I}(\mathrm{v(m,n)})+}} \end{cases} + \varepsilon(\mathrm{Gmax} - -\mathrm{G(v(m,n))}) + \Phi\mathrm{GD(v(m,n))} \qquad (4)$$

Where $_{\gamma}$, ε, Φ are the weighting terms, Imax, Gmax are the maximum intensity values of an image I and the gradient image G. $\mathrm{I}\big(\mathrm{v(m,n)}\big), \mathrm{G}\big(\mathrm{v(m,n)}\big)$ are the intensity values of the original image and gradient image at position $\mathrm{v(m,n)}$. $\overrightarrow{\mathrm{I}(\mathrm{v(m,n)})}$ is the mean intensity in the n*n square mask.

Internal energy Eint control contraction and bending and is defined as

$$\mathrm{E_{int}(v(m,n))} = \alpha(|\mathrm{v_m(m,n)}|^2 + |\mathrm{v_n(m,n)}|^2) + \beta(\ |\mathrm{v_{mm}(m,n)}|^2 + \ |\mathrm{v_{mn}(m,n)}|^2) + \ |\mathrm{v_{nn}(m,n)}|^2) \qquad (5)$$

The subscripts in the above equation are partial derivatives. The smoothness of the net is controlled using the α, β parameter

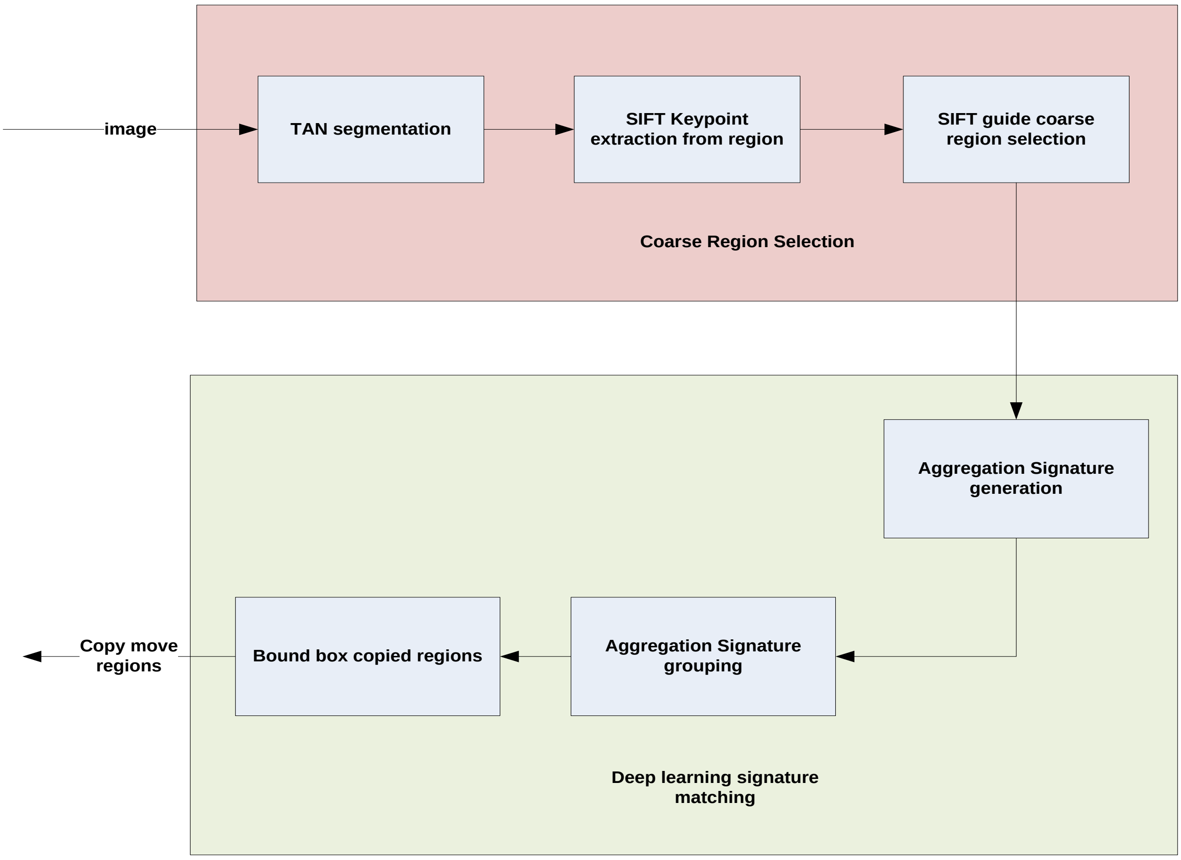
image
TAN segmentation
SIFT Keypoint extraction from region
SIFT guide coarse region selection
Coarse Region Selection
Aggregation Signature generation
Aggregation Signature grouping
Bound box copied regions
Copy move regions
Deep learning signature matching

Fig. 1 Proposed architecture

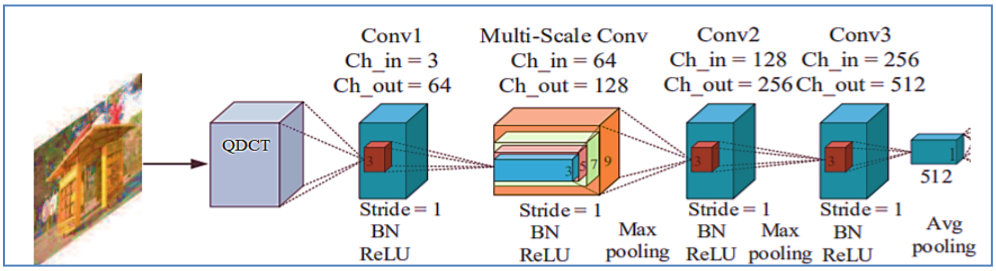

Fig. 2 Deep learning feature extraction

$$E_{ext}(v(m,n)) = \omega f[I(v(m,n))] + \frac{\rho}{|N(m,n)|}\sum_{p\varepsilon N(m,n)} \frac{1}{||v(m,n)-v(p)||} f[I(v(p))] \quad (6)$$

Where ω,ρ are the weights, $I(v(m,n))$ is the value of the intensity of the pixel at position $v(m,n)$. $N(m,n)$ is the neighborhood of node at (m,n). Function f is defined as

$$f[I(v(m,n))] = \left\{ Imax \frac{\gamma \overrightarrow{I(v(m,n))}}{-\overrightarrow{I(v(m,n))+}} + \varepsilon(Gmax - \right. \\ -G(v(m,n)))+\Phi GD(v(m,n)) \quad (7)$$

Where γ, ε, Φ are the weighting terms, $Imax, Gmax$ are the maximum intensity values of the image I and the gradient image G. $I(v(m,n)), G(v(m,n))$ are the intensity values of the original image and gradient image at position $v(m,n)$. $\overrightarrow{I(v(m,n))}$ is the mean intensity in the n*n square mask.

Internal energy Eint control contraction and bending and is defined as

$$E_{int}(v(m,n)) = \alpha(|v_m(m,n)|^2+|v_n(m,n)|^2) + \beta(|v_{mm}(m,n)|^2 + |v_{mn}(m,n)|^2) + |v_{nn}(m,n)|^2) \quad (8)$$

The subscripts in the above equation are partial derivatives. The smoothness of the net is controlled using the α, β parameter.

The energy ratio (r(n)) is calculated for all nodes, and the highest value for r(n) is selected as a threshold. A hole is opened if the r(n) of the node is greater than the threshold. Neighbours are analysed from the node, and links are deleted from the mesh based on the threshold.

However, TAN results in a large number of segments and computing deep learning aggregation signature for each segment and matching is a computationally intensive operation. TAN model results are filtered in this work based on Scale-invariant feature transform (SIFT) features [36]. SIFT is used to get all candidate key points and corresponding descriptors. For any segment X, Y got from TAN, if the neighborhood between key points in segment X to keypoint to another segment Y is calculated using the nearest neighbor distance ratio (NNDR)[30]. If the NNDR is less than a threshold, T, X and Y regions have certain similarities and must be inspected for copy-move forgery. This process is repeated for all the segments to select the coarse regions. The pseudo code of the algorithm for coarse region selection is given below

```
Algorithm  : selectCoarseRegion
Input      : Image , T
Output     : regions
 Regions ← Segment with TAN [31] (Image)
For i=1: no of regions
  Region.keypoints = SIFT(region,20);
End
Coarse_set=[]
 For i=1: no of regions
    For j=1:no of regions
         P_match←0

Num←Match_keypoint(Region(i).keypoints,
Region(j).keypoints,T)
         If Num>P_match>20
              Coarse_set.add(Regions(i));
              Coarse_set.add(Regions(j));
         End
    End
End
return  Coarse_set;
```

3.2. Deep Learning Signature Matching

A deep learning signature is generated for each coarse region returned by the region selection algorithm given in section 3.1. A deep learning aggregation signature is created, accommodating the occlusions. The aggregation signature of the image patch is formed from a set of probable noise-occluded image patches using the frequency domain deep learning model. On the image patch of the coarse region, occlusion patches of various probabilistic distributions are added. QDCT is applied to the Noised coarse region to get the low and high-frequency components. QDCT for an image $f(x,y)$ is calculated as

$$f(x,y) = A_n^q f(x,y) + \sum_{s=1}^{n}[D_{s,1}^q f(x,y) + D_{s,2}^q f(x,y) + D_{s,3}^q f(x,y)] \quad (9)$$

Where $A_n^q f(x,y)$ is the low-frequency band and $D_{s,1}^q f(x,y)$ is the high-frequency band of the image. After

QDCT is applied to the image, a low-frequency part and n groups of high-frequency parts are obtained. Low-frequency subbands are fused by averaging to reduce the dimension of the coefficients. The maximum value fusion rule is followed for high-frequency sub-bands. The low-frequency bands are fused by comparing the average of coefficients between two inputs. High-frequency bands are fused by taking the maximum coefficient value between two inputs.

The QDCT coefficients are given as input to a frequency domain convolutional neural network (Figure 3). The coefficients pass through a sequence of ReLU and max pooling layer and a final average pooling layer to provide an output of 1× 512-dimension feature vector. The CNN configuration used for feature extraction is given in Table 1. An aggregation signature is constructed from the feature vectors belonging to the same image patch as below.
A unit random vector of dimension d (d<512) is generated $\{r_0, r_1, \dots r_d\}$. The elements are selected from Gaussian distribution with 0 mean and variance as 1. The d vector is put together into a matrix D of dimension 512 × d. This is generated on time when collecting the video as input for tracking.

An inner product between the feature vectors v and the matrix D is done to get vector $u = D^T v$

For every vector u, the following transformation function tf is applied to produce the transformed feature vector $\bar{u}$.

$$tf(u) = \begin{cases} 1\ r.u\ \geq 0 \\ 0, r.u < 0 \end{cases}$$

$$\bar{u} = \{tf_{r1}(u), tf_{r2}(u), \dots . tf_{rd}(u)\}. \qquad (10)$$

The feature vectors belonging to the same image patch are now represented as a bit stream of length d called the aggregation signature of the target image patch.

Converting the features of the same patch to a binary bit stream of aggregation signature has two benefits: compressed form and reduced time complexity for matching the aggregation signature.

Once the aggregation signature is computed for all the coarse regions, the regions are grouped based on the similarity of aggregation signatures using the K-mean clustering algorithm with hamming distance for distance computation instead of typical Euclidean distance matching. Each cluster is a copied region. A bounding box is drawn in different colors for each cluster. The overall algorithm flow for deep learning signature matching is given in Figure 3.

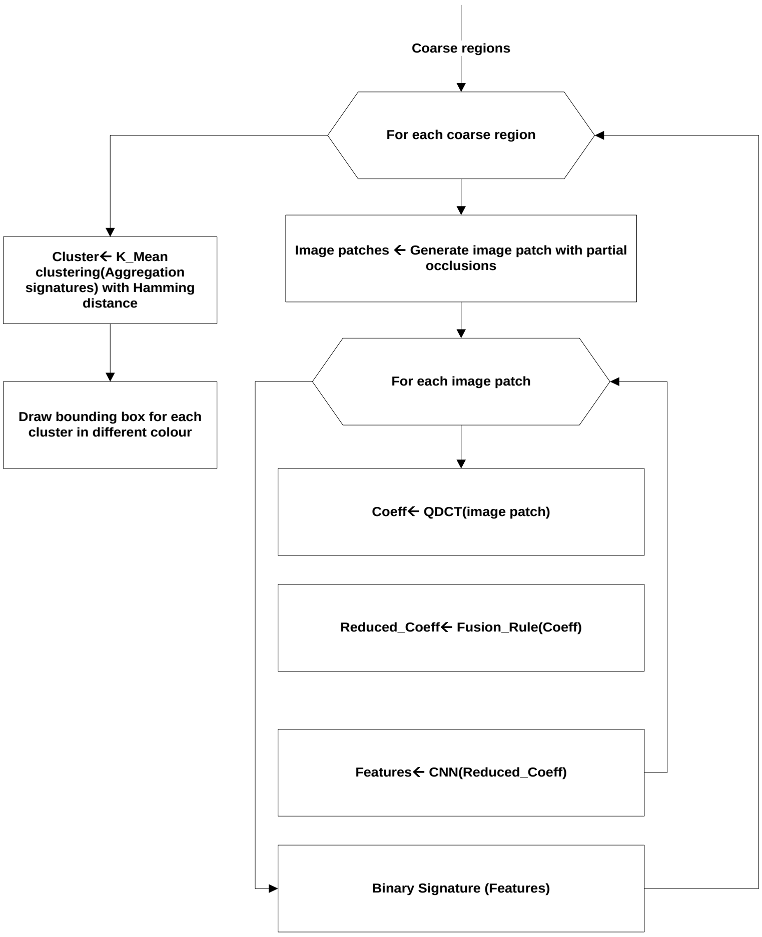
Coarse regions
For each coarse region
Cluster← K_Mean clustering(Aggregation signatures) with Hamming distance
Image patches ← Generate image patch with partial occlusions
Draw bounding box for each cluster in different colour
For each image patch
Coeff← QDCT(image patch)
Reduced_Coeff← Fusion_Rule(Coeff)
Features← CNN(Reduced_Coeff)
Binary Signature (Features)

Fig. 3 Deep learning signature matching

Table 1. Results for CMH1-4 dataset

Dataset: CMH 1-4. Total images: 108					
Method	Recall	Precision	FPR	F1	Execution time(ms)
No occlusion					
Proposed	98.21	98.77	1.98	98.63	264.31
Huang et al. (2019)	96.50	97.66	2.31	97.08	261.29
Al-Moadhen et al (2020)	97.30	98.12	2.20	98.10	292.15
Ortega et al. (2021).	98.17	97.64	2.0	98.42	294.12
Occlusion % = 20					
Proposed	97.21	97.27	2.13	96.64	264.21
Huang et al. (2019)	94.50	94.62	3.32	94.18	261.50
Al-Moadhen et al (2020)	94.30	95.22	3.22	95.12	292.51
Ortega et al. (2021).	95.17	94.62	3.01	95.47	294.37
Occlusion % = 40					
Proposed	96.21	93.77	2.32	94.71	264.31
Huang et al. (2019)	86.50	89.63	3.43	89.18	261.29
Al-Moadhen et al (2020)	89.30	90.14	3.74	89.30	292.15
Ortega et al. (2021).	90.17	89.62	3.50	89.62	294.12
Occlusion % = 60					
Proposed	93.21	92.77	2.42	94.63	263.31
Huang et al. (2019)	82.50	81.66	4.41	84.08	262.29
Al-Moadhen et al (2020)	84.30	82.12	4.56	85.10	294.15
Ortega et al. (2021).	85.17	82.64	4.25	86.42	296.12

4. Results and Discussion

The performance of the proposed solution is tested against CMH1-4 datasets [33]. CMH-1 dataset has 23 copy-move images with scaling-based forgery. CMH-2 dataset has 25 images that were copied and rotated. CMH-3 dataset has 26 images that were copied and resized. CMH-4 dataset has 34 images that were copied and then rotated and resized alternatively. The dataset images are alternated by introducing occlusions in various sizes of 20% to 60% of the original object. Standard performance metrics of precision, recall, false positive ratio, F1-score, and execution time is measured as in [34] and used for performance comparison in this work. The performance is measured for different datasets by varying the occlusion percentage. The performance is compared against the superpixel segmentation method [34] by Huang et al. (2019), Deep learning of wavelet decomposed [25] by Al-Moadhen et al. (2020) and deep learning-based detection [26] by Ortega et al. (2021).

The results for various occlusion percentage for the CMH-1 dataset is given in Table 1. The precision (Figure 4) is, on average, 1% higher than existing solutions for no occlusion images, 2% higher for occlusion 20% and 5% higher for occlusion 40% and 9% higher for occlusion 60%. Even at 60% occlusion, the precision is 92% in the proposed solution and drops about 5% from the no occlusion case. Still, other solutions experience more than a 10% drop in precision from the no occlusion case.

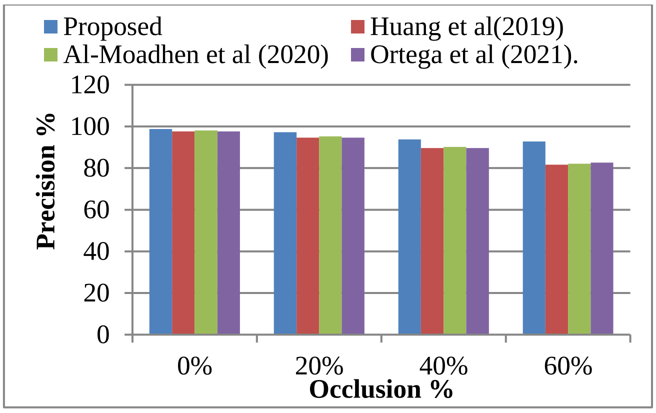

Fig. 4 Precision vs occlusion %

The probability of occlusion distribution in the construction of deep learning signature has maintained higher accuracy in the proposed solution. The equivalent deep learning-based detection method proposed by Al-moadhen et al. [25] and Ortega et al. [26] has not considered the effect of occlusions in copy-move regions.

Almost the same behavior is observed for the Recall and F1-score. As the occlusion percentage increases, FPR increases (Figure 5), but it increases only by 0.44% for occlusion, increasing from 0 to 60% in the proposed solution. Still, it is 2.1% in the case of Huang et al. (2019), 1.34% in the case of Al-Moadhen et al. (2020) and 2.25% in the case of Ortega et al. (2021).

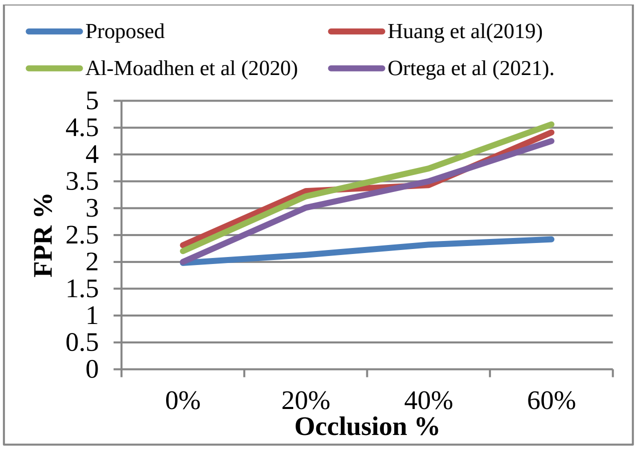

Fig. 5 FPR vs occlusion %

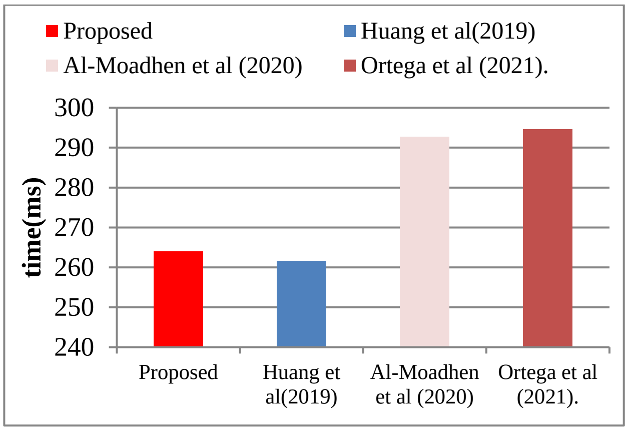

Fig. 6 Comparison of execution time

The false positive increment is lower in the proposed solution due to deformable model-based filtering and Deep learning signature matching. While other solutions have adopted shape-based features, they often mistake similar structures in the image.

The average execution time in each solution is given in Figure 6. Due to the deep learning signature, the execution time is 1% higher than Huang et al. (2019). Still, it is 10.8% higher compared to Al-Moadhen et al. (2020) and 11.6% higher compared to Ortega et al. (2021) because low complexity deep learning architecture on the frequency domain is used in the proposed solution compared to more number convolution layers used in Al-Moadhen et al. (2020) and Ortega et al. (2021). Also, the signature matching is done using hamming distance in the proposed solution, which is of low complexity to deep learning matching of each combination of images adopted in existing works.

Half total error rate is calculated as

$$HTER = \frac{\frac{FP}{FP+TN}+\frac{FN}{FN+TP}}{2} \qquad (11)$$

In the above equation, FP is a false positive. FN is a false negative, TN is a true negative, and TP is truly positive. The results for HTER across the solutions are given in Table 2.

Table 2. Comparison of HTER

Solution	HTER
Proposed	18.21
Huang et al. (2019)	20.84
Al-Moadhen et al (2020)	21.50
Ortega et al. (2021).	23.20

The proposed solution has lower HTER compared to existing works. The HTER has lowered in the proposed solution due to considering various combinations of occlusions in the signature construction. Due to this, even if there is a partial match, the fake portions are detected accurately, reducing the error.

The deep learning features for aggregation signature are computed with the QDCT method adopted in the proposed solution to other deep learning models of Densenet, VGG16 and Resnet. The accuracy results across the deep learning models are given in Table 3.

The proposed QDCT model provided at least 5% higher accuracy than other models. Accuracy has improved in the QDCT-based CNN model compared to others; the QDCT coefficients provided scope for more intricate learning compared to image-based feature extraction used in Densenet, VGG16 and Resnet models.

The ROC plot comparing the solutions is given in Figure 7.

The ROC is 0.907 in the proposed solution, which is higher than the existing works. The higher ROC indicates better sensitivity in the proposed solution. The sensitivity has increased due to considering different occlusion combinations in the signature construction process in the proposed solution. But existing works considered only scaling.

Table 3. Comparison of deep learning models

Solution	Accuracy
Proposed QDCT-based model	95
Densenet	90
VGG16	89
Resnet	87

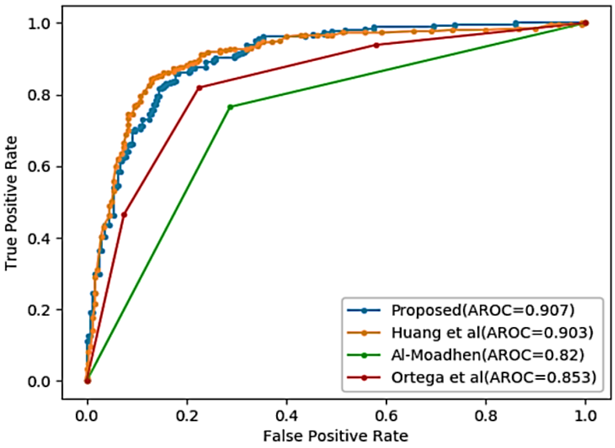

Fig. 7 Comparison of ROC

5. Conclusion

This work is a deep learning signature matching-based method for copy move forgery detection. In this two-stage solution, coarse regions for forgery detection are first selected using deformable model segmentation and SIFT-based selection. The selected coarse regions are then matched using deep learning signatures considering the probable partial occlusions. The method detected copy-move regions with more than 92%, even in the presence of more than 60% occlusions. Also, a 60% increase in occlusion caused only a 0.44% increase in FPR compared to more than 2% in the existing solution. Extending the proposed solution for colouration-based forgery is in the scope of future work.

References

[1] Mallonee, L, " Infamously Altered Photos, Before and After Their Edits," *Wired*, 2015.

[2] Wikipedia Contributors, " List of Photo Manipulation Controversies," 2019. Wikipedia.

[3] Allbeson, T.; Allan, S, " The War of Images in the Age of Trump," in Trump's Media War; Happer, C., Hoskins, A., Merrin, W., Eds.; *Springer International Publishing*: Berlin/Heidelberg, Germany, 2019.

[4] Westerlund, Mika, " The Emergence of Deepfake Technology," *A Review. Technology Innovation Management Review*, vol. 9, pp. 39-52, 2019.10.22215/Timreview/1282.

[5] Islam, M.M.; Karmakar, G.; Kamruzzaman, J.; Murshed, M. A , "Robust Forgery Detection Method for Copy–Move and Splicing Attacks in Images," *Electronics*, vol. 9, pp. 1500, 2020.

[6] R. Cristin, J. P. Ananth and V. Cyril Raj, "Illumination-Based Texture Descriptor and Fruitfly Support Vector Neural Network for Image Counterfeit Detection in Face Images," *In IET Image Processing*, vol. 12, no. 8, pp. 1439-1449, 2018.

[7] Y. Guo, X. Cao, W. Zhang and R. Wang, "Fake Colorized Image Detection," *In IEEE Transactions on Information Forensics and Security*, vol. 13, no. 8, pp. 1932-1944, 2018.

[8] H. Li, W. Luo, X. Qiu and J. Huang, " Image Counterfeit Localization Via Integrating Tampering Possibility Maps," *In IEEE Transactions on Information Forensics and Security*, vol. 12, no. 5, pp. 1240-1252, 2017.

[9] J. H. Bappy, C. Simons, L. Nataraj, B. S. Manjunath and A. K. Roy-Chowdhury, "Hybrid LSTM and Encoder-Decoder Architecture For Detection of Image Forgeries," *In IEEE Transactions on Image Processing*, vol. 28, no. 7, pp. 3286-3300, 2019.

[10] B. Chen, M. Yu, Q. Su, H. J. Shim and Y. Shi, "Fractional Quaternion Zernike Moments For Robust Color Image Copy-Move Counterfeit Detection," *IEEE Access*, vol. 6, pp. 56637-56646, 2018.

[11] Y. Li and J. Zhou, " Fast and Effective Image Copy-Move Counterfeit Detection Via Hierarchical Feature Point Matching," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 5, pp. 1307-1322, 2019.

[12] O. Mayer and M. C. Stamm, "Accurate and Efficient Image Counterfeit Detection Using Lateral Chromatic Aberration," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 7, pp. 1762-1777, 2018.

[13] X. Bi and C.-M. Pun, " Fast Reflective Offset-Guided Searching Method For Copy-Move Forgery Detection," *Information Sciences*, vol. 418–419, pp. 531–545, 2017.

[14] X.-Y. Wang, S. Li, Y.-N. Liu, Y. Niu, H.-Y. Yang, and Z.-L Zhou, "A New Keypoint-Based Copy-Move Forgery Detection For Small Smooth Regions," *Multimedia Tools and Applications*, vol. 76, no. 22, pp. 23353–23382

[15] S. Teerakanok and T. Uehara, "Copy-Move Forgery Detection Using GLCM-Based Rotation-Invariant Feature: A Preliminary Research," in Proceedings COMPSAC,pp. 365–369, 2018.

[16] M. Emam, Q. Han, Q. Li, and H. Zhang, " A Robust Detection Algorithm for Image Copy-Move Forgery in Smooth Regions," *in Proc. ICCSS*, London, U.K, pp. 119–123, 2017.

[17] Singh, B., Sharma, D.K, " Predicting Image Credibility in Fake News Over Social Media Using A Multi-Modal Approach," *Neural Computer & Application*, 2021.

[18] Zhang, Weiguo, "A Novel Counterfeit Feature Extraction Technique For Exposing Face-Swap Images Based on Deep Learning and Error Level Analysis," *Entropy (Basel, Switzerland)*, vol. 22, no. 2, pp. 249, 2020.

[19] Ghoneim, A., Muhammad, G., Amin, S. U., & Gupta, B, " Medical Image Forgery Detection for Smart Healthcare," *IEEE Communications Magazine*, vol. 56, no. 4, pp. 33-37, 2018.

[20] Thajeel, Salam & Shakir, Ali & Rasheed, Waleed & Sulong, Ghazali, " Detection Copy-Move Forgery in Image Via Quaternion Polar Harmonic Transforms," *KSII Transactions on Internet and Information Systems*, vol. 13, pp. 4005-4025, 2019. 10.3837/Tiis.2019.08.010.

[21] Liu, Y., Q. Guan, and X. Zhao, " Copy-Move Forgery Detection Based on Convolutional Kernel Network," *Multimedia Tools and Applications*, vol. 77, no. 14, pp. 18269-18293, 2018.

[22] Chou, C.-L. and J.-C. Lee, "Copy-Move Forgery Detection Based on Local Gabor Wavelets Patterns," *in Proceedings of Springer on International Conference on Security With Intelligent Computing and Big-Data Services*, vol. 733, pp 47-56, 2018.

[23] Mahmood, T., Et Al, "A Robust Technique For Copy-Move Forgery Detection and Localization in Digital Images Via Stationary Wavelet and Discrete Cosine Transform," *Journal of Visual Communication and Image Representation*, vol. 53, pp. 202-214, 2018.

[24] Hosny, K.M., 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

[25] Al-Moadhen, Ahmed Abdulhadi Ahmed & Ridha, Mustafa & Khayeat, Ali, " Splicing Detection in Color Image Based on Deep Learning of Wavelet Decomposed Image," *AIP Conference Proceedings*, 2290. 10.1063/5.0027442.

[26] Rodriguez-Ortega, Yohanna & Ballesteros, Dora & Renza, Diego, "Copy-Move Forgery Detection (CMFD) Using Deep Learning For Image and Video Forensics," *Journal of Imaging*, 7. 59. 10.3390/Jimaging7030059.

[27] Alhussain Akoum, Samia Bahlak, Nagham Abou Daher, "Image Forgery Analyse and Detection," *SSRG International Journal of Computer Science and Engineering*, vol. 8, no. 8, pp. 8-12, 2021. Crossref, Https://Doi.Org/10.14445/23488387/IJCSE-V8I8P102

[28] D. Terzopoulos "Deformable Models," *The Visual Computer*, vol. 4, pp. 306–331, 1988.

[29] Balaji V , Ajith Kumar P, Kiren Aananth A, Gunasekar N and Ciyamala Kushbu S, "Forgery Detection in Documents," *SSRG International Journal of Electronics and Communication Engineering*, vol. 5, no. 5, pp. 15-20, 2018. Crossref, Https://Doi.Org/10.14445/23488549/IJECE-V5I5P104

[30] K. Mikolajczyk, C. Schmid, "A Performance Evaluation of Local Descriptors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1615–1630, 2005.

[31] M. Bro-Nielsen, "Active Nets and Cubes," IMM, Tech. Report, 1994

[32] A. B. Z. Abidin, H. B. A. Majid, A. B. A. Samah and H. B. Hashim, " Copy-Move Image Forgery Detection Using Deep Learning Methods: A Review," *2019 6th International Conference on Research and Innovation in Information Systems (ICRIIS*), pp. 1-6, 2019.

[33] E. Silva, T. Carvalho, A. Ferreira, A. Rocha, "Going Deeper Into Copy-Move Forgery Detection: Exploring Image Telltales Via Multi-Scale Analysis and Voting Processes," *Journal of Visual Communication and Image Representation*, vol. 29, pp. 16–32 , 2015.

[34] Huang, HY., Ciou, AJ, "Copy-Move Forgery Detection for Image Forensics Using the Superpixel Segmentation and the Helmert Transformation," *Journal on Image and Video Processing*, vol. 68, 2019.

[35] A. B. Z. Abidin, H. B. A. Majid, A. B. A. Samah and H. B. Hashim, " Copy-Move Image Forgery Detection Using Deep Learning Methods: A Review," *2019 6th International Conference on Research and Innovation in Information Systems (ICRIIS)*, pp. 1-6, 2019.

[36] D.G. Lowe, "Distinctive Image Features From Scale-Invariant Keypoint," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.