A Review on Forgery Detection in Social Media Images

Mohammad Rafi[1],Rakshitha C[2],Tejaswini Surve K S[3]

mdrafi2km@yahoo.com[1],rakshithachidanand342@gmail.com[2],tejusurve162@gmail.com[3]

Department of Computer Science and Engineering, UBDTCE

Abstract

The increasing abuse of image editing software causes the authenticity of digital images questionable. The widespread availability of online social networks(OSNs) makes them the dominant channels for transmitting forged images to report fake news. The last decade has seen lot of research advancement in the area of digital image forensics, where the investigation for possible forgeries is based on post-processing of images. Deep learning approaches have shown promising results in various image classification problems but cannot find hidden patterns in digital images, which can reliably detect image forgeries. The objective of the proposed approach is to detection the accuracy. In addition to analyze the schemes and evaluate and compare their performances in terms of a proposed set of parameters, which may be used as a standard benchmark for evaluating the efficiency of any general copy-move forgery detection technique for digital images. We further incorporate the tailored noise into a robust training framework, significantly improving the robustness of the image forgery detector. The comparison results provided by them would help a user to select the most optimal forgery detection technique, depending on the author requirements. This paper discuses various forgery detection in social media images and suggests new idea of detection.

Introduction

The widespread use of social media platforms has led to an increasing in number of manipulated and fake images being shared. The digital images act as the primary sources of evidence towards any event in legal as well as media and broadcast Industries. Medias are used as the major weapons in many criminal and court case information for media industry. Digital images form a significant part of information transmitted in regular communications as well. These digital pictures are utilized to spread data to an audience on a wide scale and consequently formulate a general opinion on a large scale. Image forgery is used to refer to the act of manipulating images to showcase false information or to hide some helpful information from the images. The motive behind such manipulations can be various factors like earning money, disseminating rumors, or making false claims. The manipulated or forged images are becoming increasingly dangerous in various fields such as removing copyright watermarks, producing fake news, and being forged evidence in court, negatively affecting not only individuals but also the whole society. A large number of methods [1]–[16] have been proposed to detect and localize image forgery, so as to ensure information authenticity. Some of these forensic techniques are designed to detect specific forms of tampering, such as splicing [2], [6], copy-move [3], [7] and inpainting [5], [9], [10], while the others are to identify more complex or compound forgeries. However, few research has been done to explicitly address the design of robust forgery detection against the lossy operations in the present online social networks (OSN) platforms. Such a topic is very important because these lossy operations can severely degrade the detection performance. As shown in Fig. 1, the state-of-the-art algorithm [1] can accurately detect the forged regions from the original forgery, but the detection performance would be severely degraded when handling the forgery transmitted through Facebook. The advent of digital image manipulation tools has exacerbated the proliferation of image forgeries, necessitating robust solutions for their detection. As the core of our foregery detection system, has exhibited remarkable performance with the training accuracy of 98% and validation accuracy of 92%. This showcases its efficiency in distinguishing authentic from tampered images. The dataset utilized in for the study comprises 12,615 images, consisting of 7491 real images and 4,123 tampered images, providing a diverse and extensive testbed for evaluation. Each image is resized to a standardized 256x256 resolution. For mitigating the negative impacts of OSNs, the most critical issue is to analyze and model the noise introduced by the OSN lossy channels. However, this is a rather difficult problem mainly because the current platforms do not disclose the process for manipulating the transmitted images. Although some existing works [17], [18] revealed part of the processes adopted by OSNs, there are still many unknown operations, e.g., for Facebook, for the enhancement filtering, the allocation mechanism of the quality level and resizing the factor even the interpolation used in resizing are all not clear. OSNs adjust their image processing pipelines and making the modeling more challenging In an untampered image, all regions should exhibit uniform compression. Deviations from this uniformity may indicate digital manipulation. The

processed images are stored as numpy arrays for subsequent. The decouplement of the OSN noises are in two components: 1) *predictable noise* and 2) *unseen noise*. It is designed to simulate the predictable loss whose modeling relies on a deep neural network (DNN) with the residual learning and an embedded differentiable JPEG layer. Apparently, it is unrealistic to build a suitable model for the unseen noise from the perspective of the signal itself. To address this difficulty, we transfer our observations from the noise perspective to the forgery detector, only focusing on the noise that may cause deterioration of the detection performance. Such a strategy naturally incubates a new algorithm to model the unseen noise by utilizing the core idea of *adversarial noise*[20]. This is essential in an imperceptible perturbation that can severely degrade the model performance.

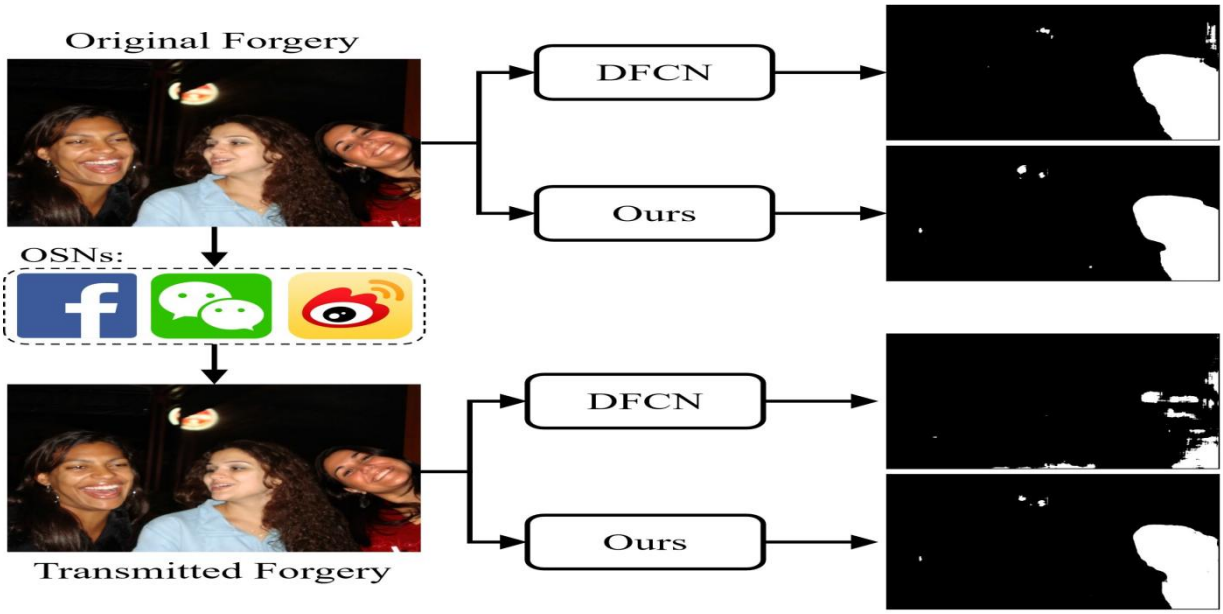
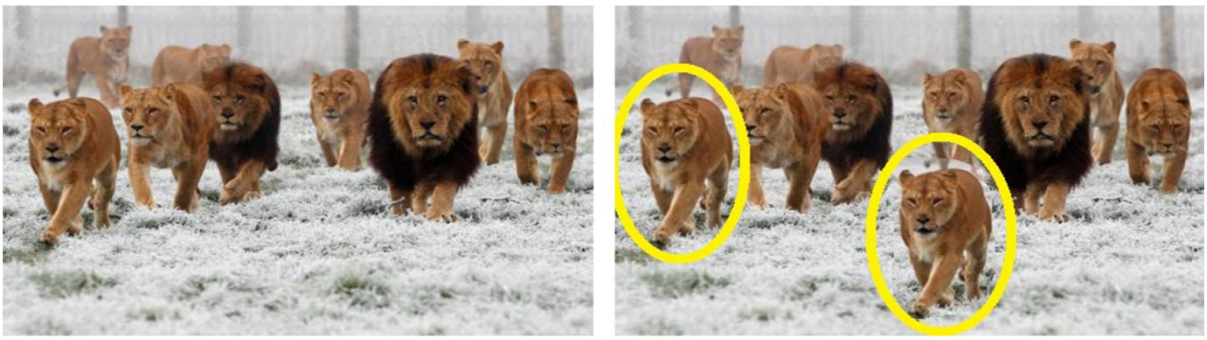


Fig. 1. The detection results of DFCN [1] and ours by using an original forgery and the forgery transmitted through OSN. The right woman in the forgery is spliced (forged).

The above figure which validates the robustness of the model against the transmission over OSN. In the design of a baseline image forgery detector, which won the top ranking in a recent certificate forgery detection competition. This baseline detector also serves as the cornerstone of the work. For the purpose of a novel training scheme for robust image forgery detection against transmission over OSNs. The training scheme will not only be the models the predictable noise involved by OSNs, but also incorporates the unseen noise through a newly proposed algorithm to further promote the robustness of the detector. Our proposed model achieves better detection performance in comparison with several state-of-the-art methods [1], [14]–[16], especially in the scenario of fighting against the transmission over OSNs. By building a public forgery dataset based on four existing datasets [21]–[24], through uploading and downloading over the platforms of Facebook, Whatsapp, Weibo, and Wechat, respectively. Since this form of digital image forgery involves duplication of regions of the same
image, the image statistics are not disturbed. This form of forgery does not lead to any significant change in the image characteristics because the texture, noise and color components do not get altered for the forged region. Rather those statistical features or characteristics remain unaltered over varied regions of the forged image, even after copy–move. Hence, to detect this form of forgery, investigation of image statistical inconsistencies is not particularly helpful. In the recent years, researchers have mainly focused on the identification of region duplication in images in order to detect copy–move forgery. To hinder duplicate regions identification in images, attackers may further modify the duplicated image regions such as by slight noise addition, blurring, rotation, re-scaling etc. Recently, the problem of identifying geometrically transformed, blurred and noise-added duplicate image regions has attracted considerable researchers interest as well. In this by presenting
the investigated of the state-of-the-art and copy-move forgery detection techniques for digital images. The operating principles of most of the state-of-the-art copy-move forgery detection techniques are ‘block-based’, which is based on the identification of duplicate image blocks. By providing the readers a detailed survey of ‘block-based’ region duplication techniques for digital images, along with an evaluation, analysis and comparison of their performance efficiencies will increase, through a three way standard platform, which would enable the readers to select a particular copy-move forgery detection scheme. The most common forms of modification attacks to digital images include image retouching [21], image splicing [22] and copy–move forgery [23–30]. In image retouching, features of an images are altered, so that the modifications are difficult to be detected. Image splicing is the form of digital image forgery where the forger combines regions from multiple images into a single image, so as to form a natural looking composite image. Such modifications are detectable by investigating inconsistencies in natural

statistics of the image [30]. In copy–move [23,24] form of attack on digital images, regions of an image are copied and pasted onto itself, at some different locations, with the malicious intention to obscure or repeat significant objects in the image. For example, Fig. 2 presents an example of a copy–move attack on an image, depicting a forest scene with lions moving in a group. The original image has been shown in Figs. 2*a* and *b* as its copy–move forged version, where a lioness object has been copied from the left most position of the original scene and pasted onto itself different location. In the forged image, one can find one additional lioness in the front.



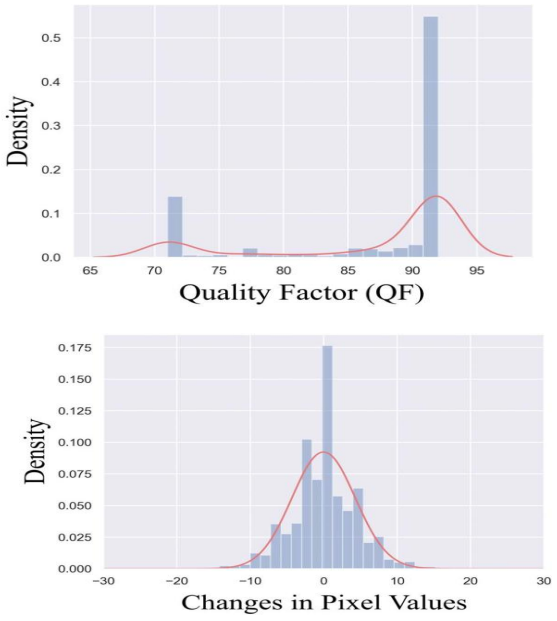
a *b*

Fig. 2 Example of copy–move forgery **(a)** Original image, **(b)** Forged image (duplicated object highlighted)

Literature Review

Image Forgery Detection:

Many forensic methods (e.g., [2]–[10] and references) have been proposed to verify the authenticity of digital images. These methods detect the forged regions through the *specific* artifacts left by the tampering operations like splicing [2,6], copy-move [3,7], median filtering inpainting [5,9,10], etc. More specifically, Lyu *et al.* [2] introduced an effective method for the splicing detection by revealing inconsistencies in local noise levels. Through solving the key point matching problems over a massive number of key points, Li and Zhou [3] developed a fast hierarchical matching strategy for the detection of copy-move forgeries. As for the forensic detection of the median filtering, Kang *et al.* [4] adopted an autoregressive model to analyze the statistical properties of the median filter residual. To extract the evidence of the inpainting forgeries, Li *et al.* [5] proposed a diffusion-based detection method by analyzing the local variance of the image Laplacian along the isophote direction. With the success of neural networks in various fields, many deep learning based approaches [6]–[10] have been developed for detecting these specific forgeries. Unfortunately, these forensic approaches can only be applied to detect specific tampering manipulations, severely limiting their practical usefulness, as the prior knowledge regarding the forgery types is usually unavailable. To better fit the p To better fit the practical requirements, in recent years, more and more methods have been developed to address the problem of detecting general types of forgeries [1], [11]–[16], among which the deep learning based methods are the most successful. Along this line of research, Wu *et al.* [14] proposed the MT-Net, a general forgery detection or localization network, which first extracts image manipulation features and then identifies anomalous regions by assessing how different a local feature is from its reference features. Mayer and Stamm recently [15] introduced the forensic similarity to determine whether two image patches contain the same or different forensic traces. From the perspective of the camera fingerprint, Cozzolino and Verdoliva designed a method for extracting a camera model fingerprint, called noiseprint, so as to disclose the forged regions via suppressing the scene contents while enhancing the model-related artifacts [16]. For learning the traces of generic forgeries, Zhuang *et. al* [1] utilized a training data generation strategy by resorting to Photoshop scripting.



Online Social Network (OSN):

The popularity of various OSN platforms, e.g., Facebook, Whatsapp, Wechat, Weibo, etc, significantly implifies the dissemination and sharing of images. As indicated by many existing works [17], [18], almost all OSNs manipulate the uploaded images in a lossy fashion. The noise introduced by these lossy operations could severely affect the effectiveness of forensic methods. By taking Facebook as an example, as discovered in the previous works [17], [18], [25], these manipulations mainly consist of four stages: format

conversion, resizing, enhancement filtering, and JPEG compression. Specifically, the uploaded image is first converted into the pixel domain, where the truncation is used to ensure the pixel values are within the existing. After that, resizing would be applied if the resolution of the image is above 2048 pixels. Subsequently, some selected blocks in the image undergo highly adaptive and complex enhancement filtering. It is very challenging to precisely know these enhancement filtering operations due to their adaptiveness. Finally, the image is subject to a round of JPEG compression with a quality factor (QF) *adaptively* determined according to the image content. Through the analysis of the dataset provided, the QF values used by Facebook range from 71 to 95, where a more detailed distribution is shown in Fig. 3a,3b how the pixel values change when an image is transmitted through Facebook. Although the image manipulations on different OSN platforms are different.

Methodology

Dimensionality reduction-based copy–move forgery detection here in this section, detailing by the operations of different algorithms belonging to the class of dimensionality reduction based copy–move forgery detection. They are: The PCA-based algorithm, The SVD-based algorithm and The PCA-DCT-based algorithm. PCA-based copy–move forgery detection is the algorithm which divides an image into overlapping blocks. Each block is sorted lexicographically with respect to the pixel intensities. Each sorted block is stored into one row of a matrix. When w × h image is divided into B × B as overlapping blocks, the covariance of each such block is computed as

$$C_m = \sum_{i=1}^{N_{\text{total}}} \boldsymbol{x}_i \boldsymbol{x}_i^{\text{T}}$$

where $\boldsymbol{x}_i$ represents a block for i = 1, 2, 3, ..., N_{total} and $N_{\text{total}} = (w - B + 1) \times (h - B + 1)$ represent the total number of blocks. The principal components of C_m are defined by the eigenvectors $\boldsymbol{e}_j$ for j = 1, 2, 3 so on, B (of C_m) corresponding to the eigenvalues λ_j (j = 1, 2, …, b and $\lambda_1 > \lambda_2 > \ldots > \lambda_B$). Each image block can be linearly represented in terms of the eigenvectors as

$$\boldsymbol{x}_i = \sum_{j=1}^{B} a_j x_j$$

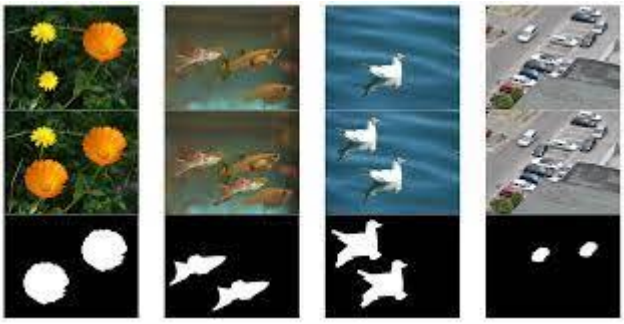
where $a_j = \boldsymbol{x}_i^{\text{T}} \boldsymbol{e}_j$ show the new representation for each image block. Each vector $\boldsymbol{x}_i$ is shortened to first N_t terms, N_t being a user defined parameter, in order to reduce the dimensionality of each block and generate a new N_t – dimensional representation of C_m. Following are the steps for detection of duplicate image blocks.

i.Matrix $\boldsymbol{S}$ is obtained by sorting row wise lexicographically.

ii. Let s_i denote the ith row of matrix $\boldsymbol{S}$. The row s_i of the matrix $\boldsymbol{S}$ is represented using the tuple $(\boldsymbol{x}_i, y_i)$, such that $(\boldsymbol{x}_i, y_i)$ represents a block's image co-ordinates.

iii. A list L is constructed that stores every pair of rows s_i, s_j such that $|i - j| < N_n$, where N_n is a user-defined parameter denoting the number of neighboring rows to be searched.

PCA-based copy–move forgery detection



iv. The offset frequency for a pair s_i, s_j, present in list L,is calculated as

$$\begin{matrix} (x_j - \boldsymbol{x}_i,\ y_j - y_i) & \text{if} & (\boldsymbol{x}_i - x_j) < 0 \\ (\boldsymbol{x}_i - x_j,\ y_i - y_j) & \text{if} & (\boldsymbol{x}_i - x_j) > 0 \\ (\boldsymbol{x}_i - x_j,\ y_i - y_j) & \text{if} & (\boldsymbol{x}_i = x_j) \end{matrix}$$

$$C_m = \sum_{i=1}^{N_{\text{total}}} \boldsymbol{x}_i \boldsymbol{x}_i^{\text{T}}$$

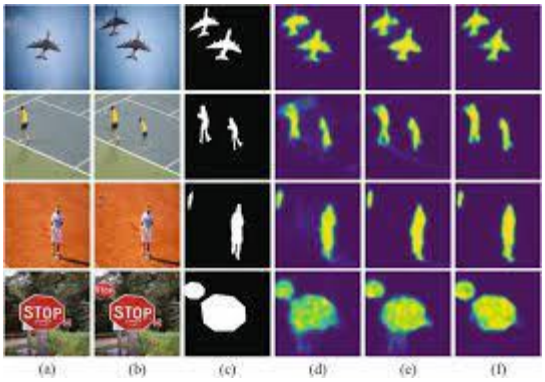
v. The offset magnitude for pair s_i, s_j in L, is calculated as

$$\sqrt{(\boldsymbol{x}_i - x_j)^2 + (y_i - y_j)^2}$$

vi. The pairs having offset frequency less than N_f and offset magnitude less than N_d are discarded. N_f and N_d denote the minimum frequency threshold and the minimum offset threshold as chosen by the user.

vii The maliciously duplicated blocks are represented by the remaining pairs of rows contained in L. The results of duplicate regions detection, using the algorithm which has been shown. In the above image, where the size of forgery has been varied as 20, 30 and 40% of the entire image. The copy–move forgery detection results using the method, for the manually forged images, which has been shown.

SVD-based copy–move forgery detection



The SVD is an algebraic transform, which finds wide range of application in several fields such as image and signal processing, pattern analysis, data compression and scientific computing. SVD decomposes one block of an image into three matrices $\boldsymbol{U}$, $\boldsymbol{S}$ and $\boldsymbol{V}$ each of which is sufficiently smaller compared with the original image block, they preserve the inherent features of the image block. So in the process of feature extraction from these matrices becomes computationally lesser intensive and also consumes lesser memory for storage compared to the original block. Utilizing singular value, the SVD technique has the extraction of unique feature vectors of image blocks which reduces dimension of block features. The steps for feature vectors extraction from image blocks [9], while reducing the dimensionality of feature space as follows:

i. A $w \times h$ image is divided into $(w - B + 1) \times (h - B + 1)$ overlapping blocks, each of size $B \times B$ pixels.

ii. Let $\boldsymbol{A}$ be a $B \times B$ matrix representing one block of the image. $\boldsymbol{A}$ is decomposed into its singular value matrices $\boldsymbol{U}$, $\boldsymbol{S}$ and $\boldsymbol{V}$, each of the dimension $B \times B$ as:

$$\boldsymbol{A} = \boldsymbol{USV}^{\mathrm{T}}$$

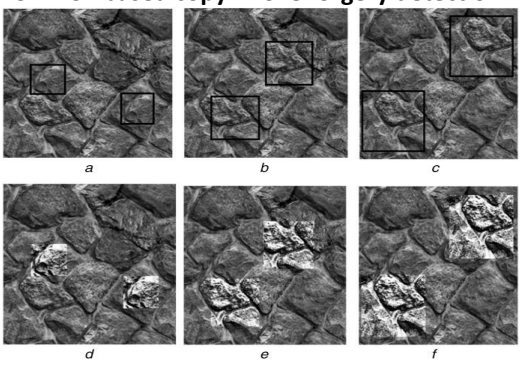
where each of $\boldsymbol{U}$, $\boldsymbol{S}$ and $\boldsymbol{V}$ are real number matrices. $\boldsymbol{S}$ is a diagonal singular value matrix of the form.

v. All the pairs of rows in the feature vector matrix, whose Euclidean distance is more than the similarity threshold Td , are discarded as they are considered to be similar blocks of the image. Further verification is performed on the remaining pairs that pass this stage of elimination.

vi. For a given pair of image blocks u and v, with the blocks image co-ordinates (i, j) and (k, l), respectively, the *Chebyshev distance* between u and v is computed as $C_{uv} = \max | \mathrm{abs}(i - k), \mathrm{abs}(j - l)|$

vii. If $C_{uv} \geq T_s$, then blocks u and v are labelled as suspected duplicate blocks, where T_s is chosen as a threshold representing minimum separation between duplicate image regions. The detection of duplicate regions using the technique.

PCA-DCT-based copy–move forgery detection



The dimensionalities of the feature vectors. In detail, the steps of the algorithm can be presented as follows, For each image block represented by $\boldsymbol{A}$, the positive diagonal entries in $\boldsymbol{S}$ are sorted in non-increasing order and stored into one row of a matrix, called the *feature vector matrix*. Each row of this matrix represents the features of one block. The Euclidean distances $D(u, v)$ between two rows, u and v, of the feature vector matrix, are computed as

$$D(u, v) = \sqrt{\sum_{i=1}^{r} (u_i - v_i)^2}$$

where $u = (u1, u2, \ldots, ur)$ and $v = (v1, v2, \ldots, vr)$.

i. A $w \times h$ image is divided into $(w - B + 1) \times (h - B + 1)$ overlapping blocks, each of size $B \times B$ pixels.

ii. Next, apply DCT on each image block and store the quantized coefficients for each block into one row of the feature matrix $\boldsymbol{M}$.

iii. From every row of the matrix, by considering only the first $\lceil q \times B_2 \rceil$ elements for further processing, where $q \in (0, 1)$ hence, with a $\{(w - B + 1)(h - B + 1)\} \times \lceil q \times B_2 \rceil$ matrix.).

iv. The dimensionality of the feature matrix $\boldsymbol{M}$ is reduced through application of PCA.

v. Next, we apply a lexicographic sorting on the rows of dimensionality reduced $\boldsymbol{M}$. The identical rows are located in the sorted matrix. The duplicate image blocks are none other than those, which corresponding to the identical pairs of rows $\boldsymbol{M}$.

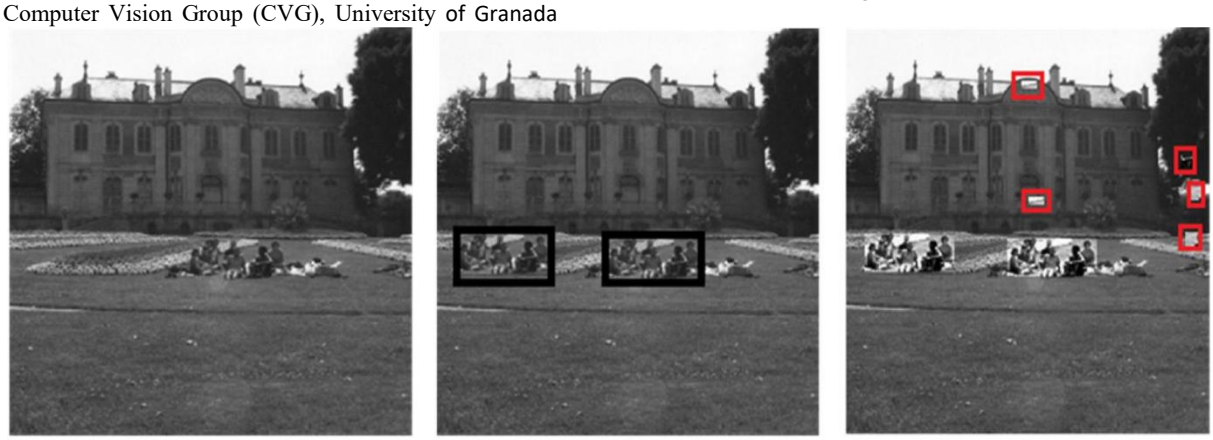
vi. The shift vector or movement vector $\boldsymbol{M}_v$ for a pair of matching blocks is calculated as $\boldsymbol{M}_v = (m_{v1}, m_{v2}) = (i_1 - j_1, i_2 - j_2)$ where (i_1, i_2) and (j_1, j_2) are the positions of two matching blocks. The movement vectors $-\boldsymbol{M}_v$ and $+\boldsymbol{M}_v$ represent the same movement. Hence, by considering the movement vectors absolute values $|\boldsymbol{M}_v|$.

vii. A matching vector counter C is used to record the frequency of occurrence of every matching block pair. Initially, all matching vector counters are set to zero. For every pair of matching blocks (vectors or rows of $\boldsymbol{M}$), the counter C is incremented by one. $C(m_{v1}, m_{v2}) = C(m_{v1}, m_{v2}) + 1$.

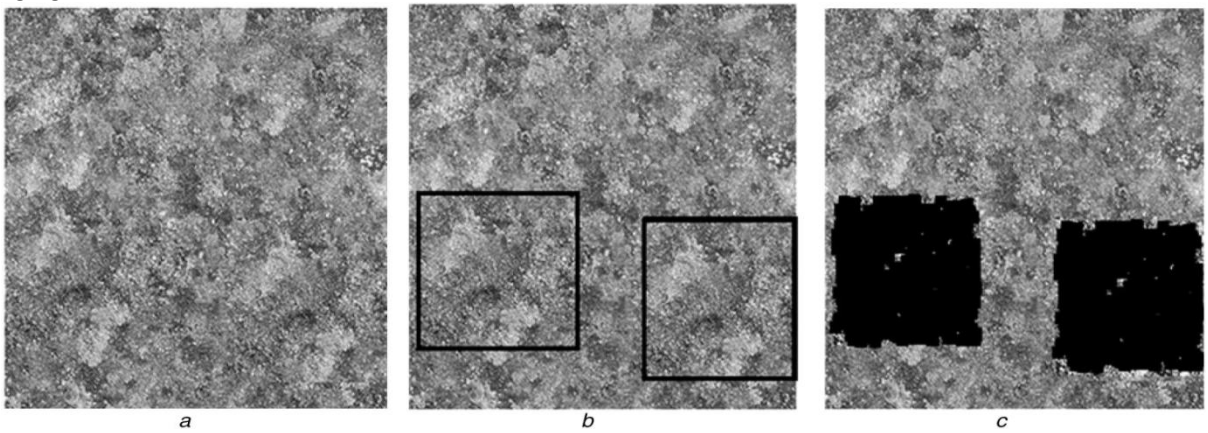
viii. The matching vector counter values are computed for all movement vectors $\boldsymbol{M}_{v1}$, $\boldsymbol{M}_{v2}$ so on. At the end of the matching process, the duplicate blocks are identified by the following criteria $C(\boldsymbol{M}_v) > T$ where T is a user-defined threshold. The block/vector pairs satisfying the above criteria are identified to be duplicates.

Experimental Results

In this section, by presenting the performance evaluation results for the block-based copy–move forgery detection algorithms discussed. All implementations are performed in MATLAB Image Processing Toolbox. Our test dataset consists of 50 standard image processing test images of size 256 × 256 pixels. By collecting these images from Computer Vision Group (CVG), University of Granada (UGR). Image Database and University of Southern California (USC), Signal and Image Processing Institute Image Database. For sake of experiments, which have manually forged our test images. The area of the duplicate region in each test image was varied from 10 to 40% of the entire image.



FPs in copy–move forgery detection using DyWT (a) Original image, (b) Forged image, (c) Output image (including FPs) highlighted



*FNs in copy–move forgery detection using DyWT (a)*Original image ***(b)*** Forged image regions duplicated ***(c)*** Detected duplicate regions darkened.

In this experiments, dividing manually forged test images into uniform overlapping blocks of size $B \times B$ pixels, where the B is varied from 6 to 36. The performance characteristics of the various techniques presented in terms of DA, have been presented in Fig. The results shown in the plots are the averages taken over all our test images. From the Fig, it is evident for all the algorithms, the DA increases with increasing forgery size. The maximum and average DA of all discussed algorithms are presented. Among all the techniques, the CWT-based method exhibits the best DA of 99.59% when the forgery size is 40%. This is due to the inherent properties of CWT exploited in the class of algorithms such as rotation invariance, robustness to noise and multi-level representation, which makes it an extremely efficient method for feature extraction. The characteristic FPR variation for region duplication detection for all the techniques versus unit block size and forgery size. It is the evident that for all the algorithms, the FPR decreases with increasing the forgery size. By wavelet-based copy–move forgery detection methods, several identical blocks get falsely detected at the boundaries of the images they contribute to the FPs which is not possible to be eliminated completely by adjusting the threshold. Presenting the false copy–move forgery detection results, for all the schemes, according to varying forgery sizes. Among all the schemes, the DCT-based techniques demonstrate the lowest rate of FPs. Similarly, the results for region duplication attacks falsely missed by the state-of-the-art techniques have been presented. Figure shows the plot of FNR versus unit block size and forgery size. From the Fig, it is evident that the FN detection rate diminishes with increasing forgery size, for all the techniques. It may be by the observation that the FN detection rate is trivial for all the techniques presented in this paper. The major challenge in this area of forensic research is to minimize the rate of FPs, as is evident. From the Figures, it may be observed that for any copy–

move forgery detection technique, its DA and FN forgery detection characteristics are inversely proportional to each other. This is due to the fact that DA is directly computed depending on the number of correctly detected copy–moved pixels, while the FNR is determined by the number of undetected copy–moved pixels. The computational complexity of any block-based copy–move forgery technique increases as the unit detection block size is reduced. On the other hand, a smaller unit detection block size ensures higher DA. Hence, in such algorithms, it is desirable to obtain a correct trade-off between DA and computational complexity, by selecting an appropriate unit block size. In this regards, the experimental results may help a user to select the most suitable method and unit block size, according to the requirements.

Conclusion

In the last decade, there has been quite a lot of researches in the direction of image forgery detection. Specifically in the field of copy–move forgery or region duplication detection in images has gained a lot of research interest due to the fact that this form of forgery is one of the most primitive forms of attacks on digital images. However, it is not trivial to detect this form of forgery because the natural statistical properties of the images are not altered here. In this paper, by providing a detailed review of state-of-the-art copy–move forgery detection algorithms, their implementation, performance evaluation and comparison. In this paper, by introducing a set of standard parameters with respect to which by having performed the experiments for the performance evaluation and comparison. The parameters introduced in this paper encompass three different dimensions of conventional forgery detection operations. The proposed parameterization would help the users select an appropriate forgery detection algorithm according to the requirements, and the expected forgery type. Future research in this direction would include incorporating more parameters into the proposed platform in order to optimize its efficiency in terms of image forgery detection evaluation and comparison. By proposing a novel training scheme for improving the robustness of the image forgery detection against various OSN based transmissions. The proposed scheme is designed with the assistance of the modeling of a predictable noise $\boldsymbol{\tau}$ as well as an intentionally introduced unseen noise. Experimental results are provided to demonstrate the superiority of our scheme compared with several state-of-the-art methods. Further, by building an OSN transmitted forgery dataset for future research of the forensic community. As the future work, may be by extending the proposed robust training scheme to deal with more complex. degradation scenarios, such as screen capturing, printing and re-photographing, etc. Additionally, investigating whether an image restoration network can be used to assist the forgery detection in severely degraded scenarios.

Robustness evaluations

Although the proposed scheme is mainly designed to counter the lossy operations conducted by OSNs, and also like to evaluate its robustness under some more commonly used degradation scenarios, such as noise addition,cropping, resizing, blurring, and standalone JPEG compression. Such evaluation is very critical in real-world cases because these types of post-processing operations are often adopted to erase or conceal the forged artifacts. To this end, applying these post-processing operations to the original test set **Columbia** and report the quantitative comparisons. For the convenience of demonstration, utilizing a unified parameter p for controlling the magnitudes of different operations. The origin of the horizontal axis ($p = 0$) corresponds to the case without any post-processing. It can be observed, the competitors [12,27] cannot perform consistently with the increase of the perturbation intensity, while this method can generalize well to defeat these post processing operations.

References

[1] P. Zhuang, H. Li, S. Tan, B. Li, and J. Huang, “Image tampering localization using a dense fully convolutional network,” IEEE Trans. Inf. Forensics Security, vol. 16, pp. 2986–2999, 2021.

[2] S. Lyu, X. Pan, and X. Zhang, “Exposing region splicing forgeries with blind local noise estimation,” Int. J. Comput. Vis., vol. 110, no. 2, pp. 202–221, Nov. 2014.

[3] 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.

[4] X. Kang, M. C. Stamm, A. Peng, and K. J. R. Liu, “Robust median filtering forensics using an autoregressive model,” IEEE Trans. Inf. Forensics Security, vol. 8, no. 9, pp. 1456–1468, Sep. 2013.

[5] H. Li, W. Luo, and J. Huang, “Localization of diffusion-based inpainting in digital images,” IEEE Trans. Inf. Forensics Security, vol. 12, no. 12, pp. 3050–3064, Dec. 2017.

[6] M. Huh, A. Liu, A. Owens, and A. A. Efros, “Fighting fake news: Image splice detection via learned self-consistency,” in Proc. Eur. Conf. Comput. Vis., 2018, pp. 101–117.

[7] J.-L. Zhong and C.-M. Pun, “An end-to-end dense-InceptionNet for image copy-move forgery detection,” IEEE Trans. Inf. Forensics Security, vol. 15, pp. 2134–2146, 2020.

[8] J. Chen, X. Kang, Y. Liu, and Z. J. Wang, "Median filtering forensics based on convolutional neural networks," IEEE Signal Process. Lett., vol. 22, no. 11, pp. 1849–1853, Nov. 2015.

[9] H. Wu and J. Zhou, "IID-Net: Image inpainting detection network via neural architecture search and attention," IEEE Trans. Circuits Syst.Video Technol., early access, Apr. 22, 2021, doi: 10.1109/TCSVT. 2021.3075039.

[10] A. Li et al., "Noise doesn't lie: Towards universal detection of deep inpainting," in Proc. 13th Int. Joint Conf. Artif. Intell., Aug. 2021, pp. 1–7.

[11] D. Cozzolino, G. Poggi, and L. Verdoliva, "Splicebuster: A new blind image splicing detector," in Proc. IEEE Int. Workshop Inf. Forensics Secur. (WIFS), Nov. 2015, pp. 1–6.

[12] L. Bondi, S. Lameri, D. Guera, P. Bestagini, E. J. Delp, and S. Tubaro, "Tampering detection and localization through clustering of camera based CNN features," in Proc. IEEE Conf. Comput. Vis. Pattern Recog nit. Workshops (CVPRW), Jul. 2017, pp. 1855–1864.

[13] B. Bayar and M. C. Stamm, "Constrained convolutional neural networks: A new approach towards general purpose image manipulation detection," IEEE Trans. Inf. Forensics Security, vol. 13, no. 11, pp. 2691–2706, Nov. 2018.

[14] Y. Wu, W. Abdalmageed, and P. Natarajan, "ManTra-Net: Manipulation tracing network for detection and localization of image forgeries with anomalous features," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2019, pp. 9543–9552.

[15] O. Mayer and M. C. Stamm, "Forensic similarity for digital images," IEEE Trans. Inf. Forensics Security, vol. 15, pp. 1331–1346, 2020.

[16] D. Cozzolino and L. Verdoliva, "Noiseprint: A CNN-based camera model fingerprint," IEEE Trans. Inf. Forensics Security, vol. 15, pp. 114–159, 2020.

[17] W. Sun, J. Zhou, R. Lyu, and S. Zhu, "Processing-aware privacy preserving photo sharing over online social networks," in Proc. 24th ACM Int. Conf. Multimedia, Oct. 2016, pp. 581–585.

[18] W. Sun, J. Zhou, Y. Li, M. Cheung, and J. She, "Robust high-capacity watermarking over online social network shared images," IEEE Trans. Circuits Syst. Video Technol., vol. 31, no. 3, pp. 1208–1221, Mar. 2021.

[19] Security Ai Competition: Forgery Detection on Certificate Image. Accessed: Jan. 23, 2022. [Online]. Available:https://tianchi.aliyun.com/competition/entrance /531812/information

[20] C. Szegedy et al., "Intriguing properties of neural networks," in Proc. Int. Conf. Learn. Representat., 2014, pp. 1–10.

[21]Savchenko, V., Kojekine, N., Unno, H.: 'A practical image retouching method'. Proc. First Int. Symp. Cyber Worlds, 2002, pp. 480–487

[22] Redi, J.A., Taktak, W., Dugelay, J.: 'Image splicing detection using 2D phase congruency and statistical moments of characteristic function'. Society of Photo-optical Instrumentation Engineers (SPIE) Conf. Series, 2007, vol. **6505**, p. 26

[23] Fridrich, A.J., Soukal, B.D., Lukáš, A.J.: 'Detection of copy–move forgery in digital images'. Proc. Digital Forensic Research Workshop, 2003

[24] Farid, A.P., Popescu, A.C.: 'Exposing digital forgeries by detecting duplicated image region'. Technical Report, Hanover, Department of Computer Science, Dartmouth College, USA, 2004

[25] Kang, X., Wei, S.: 'Identifying tampered regions using singular value decomposition in digital image forensics'. Int. Conf. Computer Science and Software Engineering, 2009, vol. **3**, pp. 926–930

[26] Zhang, J., Feng, Z., Su, Y.: 'A new approach for detecting copy–move forgery in digital images'. 11th IEEE Singapore Int. Conf. Communication Systems, 2008, pp. 362–366

[27] Muhammad, G., Hussain, M., Bebisi, G.: 'Passive copy–move image forgery detection using undecimated dyadic wavelet transform', Digit. Invest., 2012, **9**, (1), pp. 49–57

[28] Yang, J., Ran, P., Tan, J.: 'Digital image forgery forensics by using undecimated dyadic wavelet transform and Zernike moments', J. Comput. Inf. Syst., 2013, **9**, (16), pp. 6399–6408

[29] Bayram, S., Sencar, H.T., Memon, T.N.: 'An efficient and robust method for detecting copy–move forgery'. IEEE Int. Conf. Acoustics, Speech and Signal Processing, 2009, pp. 1053–1056

[30] Huang, Y., Lu, W., Sun, W., et al.: 'Improved DCT-based detection of copy–move forgery in images', Forensic Sci. Int., 2011, **206**, (1), pp. 178–184