 Forgery localization in images based on joint statistics of image blocks with neighbouring blocks

Divya Singhal | Abhinav Gupta

Abstract
Flawless image forensic analysis necessitates precise identification of tampering regions in digital images along with the determination of state of an image (original or forged). Most of the efforts towards localization of forgeries involve localization of information-changing forgeries with less focus towards localization of information-preserving forgeries. This paper proposes a new information-preserving forgery localization method to localize 10 different tampering types exploiting the fact that joint statistics of locally forged regions and neighbouring original regions is disturbed. The proposed 18-dimensional detector is trained using ensemble classifier for fine-grain identification of forged regions in post-JPEG compressed images and results are compared with two recent state-of-the-art detectors. The results demonstrate the effectiveness of the proposed detector with significant performance gain over existing detectors, particularly for localization in images that are compressed with low quality factors of compression.

1 | INTRODUCTION

With high vulnerability of digital images to be doctored, various image forensic methods are developed in last decade to examine the authenticity of digital images in a passive way. Since, the doctored images are created using various manipulation types, researchers aim to develop forensic methods to uncover different information-changing and information-preserving manipulations. Information-changing manipulations include the forgery types that modify information associated with the original images, for example copy-move forgery and splicing. On the other hand, information-preserving manipulations do not visually alter images, but are used to destroy the identifiable fingerprints of underlying information-changing manipulations to deceive existing forensic methods. Additionally, such manipulations can be used to create more realistic forged images. In prior works, information-preserving manipulations are used as anti-forensic methods for underlying tampering types [1, 2]. Commonly used information-preserving manipulations include image filtering, noise addition and JPEG compression.

In recent years, various forensic methods have been developed for detection of both information-changing and information-preserving manipulation types, that is copy-move forgery [3], splicing [4], median filtering [5–9], JPEG compression [10], contrast enhancement [11, 12] and re-sampling [13]. Such methods are known as dedicated forensic methods that can be used to detect particular forgery types. However, generation of visually realistic forged images persuade the forgers to use a sequence of image processing operations. A dedicated forensic method designed to detect a specific forgery type may not be able to disclose the processing steps of a targeted image. Additionally, in the blind image authentication scenario, it is difficult to identify the suitable use cases of dedicated detectors. Therefore, in order to overcome the limitations of dedicated forensic methods, the focus has been shifted towards the development of generalized image forensic methods [14–22] to deal with multiple image processing operations.

Typically, practical forensic scenarios encounter doctored images that are created to incorrectly communicate the information associated with the images. As a general practice, the development of such images uses a careful mix of locally applied (forgeries applied to specific regions in images) information-changing and information-preserving forgeries. For example, a part of an image can be re-scaled and pasted onto another image to create a spliced image. In this case, information-preserving
manipulation (re-scaling) is used to make a convincing match between the size of the object being copied from an image and the image to which it is pasted. The semantics and forgery locations in such forged images can be determined by classifying small regions of an image as pristine or tampered. Post-processing operations, image spatial domain filtering methods, resizing, rotation and JPEG compression and so forth, are aimed to disguise most obvious artefacts of copy-move and splicing forgeries. In practice, forgeries are applied only in particular regions of an image rather than tampering the complete image. In literature, such forgery is called local forgery in images. Based on the above facts, the importance of general-purpose image forensics and tampering localization in the field of image forensics can be understood.

We are interested in exposing the forged regions of images, tampered with various information-preserving manipulation types. The binary classification label is generated for each non-overlapping image block to localize the forged region. A novel detector of 18 dimensions, based on relative statistics of an image block with its neighbouring blocks, is proposed for classification of image blocks. Unlike state-of-the-art detectors, the proposed work is employing the joint statistics of image neighbouring blocks for the first time to design a localization detector. However, existing detectors have used the statistics of independent image blocks for forgery localization. The proposed detector is used to localize 10 different information-preserving operations in images. These 10 manipulation types include average filtering, median filtering, contrast enhancement, Gaussian filtering, Wiener filtering, JPEG compression, JPEG 2000 compression, re-scaling, white noise addition and unsharp masking. The main contributions of proposed work as compared to state-of-the-art methods are: (1) development of general-purpose localization detector that can identify 10 different tampering operation types using a single framework, (2) identification of location of forged regions in images when forgery artefacts are manipulated using post-JPEG compression and (3) relatively smaller feature set dimensions (18-dimensions) than state-of-the-art general-purpose detectors [21, 22].

The rest of the paper is structured as follows. In Section 2, related work done in the field of general-purpose image forensics is discussed. Section 3 describes the motivation for the proposed work and relative statistics of image blocks with neighbouring blocks. Section 4 discusses the novel feature set followed by feature set analysis in Section 5. The databases and the experimental method are discussed in Section 6. Results are discussed in Sections 7 and Section 8 concludes the work.

2 LITERATURE REVIEW

Ineffectiveness of dedicated image forensics for blind image verification motivated the researchers to develop general-purpose image forensics. Additionally, localization of information-preserving forgeries in images unfolds the suspected image regions. In recent years, the field of general-purpose image forensics has emerged to develop forensic methods for both full sized images and small patches in an image. In [14], a manipulation detection framework for image patches of size $512 \times 512$ is developed by fusing multiple image forensic detectors into an ensemble detector. The multiple forensic detectors are developed by analyzing the derivative correlation features (1536 dimensions) introduced by de-mosaicing algorithm. The detector is able to localize 13 different manipulation types in high resolution images with low probability of error. However, localization performance of the detector for small sized (down to $32 \times 32$) image patches and for post-JPEG compressed images has not been analyzed. Another general-purpose detector [13] is developed by extracting residual domain image blocks and computing co-occurrence histograms of quantized image residual blocks. Although, the detector can effectively localize four image processing operations in uncompressed images for image block size $128 \times 128$ but localization performance in compressed images has not been analyzed.

Further, in [16], a general-purpose detector is designed using already defined steganalysis based feature sets, subtractive pixel adjacency matrix (SPAM) [23], spatial rich model (SRM) [24] and local binary patterns (LBP) [25]. However, high dimensionality of feature sets SPAM (686-dimensions), SRM (3461-dimensions) and LBP (22, 153-dimensions) make simulation of feature sets computationally demanding. The selected features are employed for identification of five different image processing operations. The method does not address the problem of manipulation localization in images. In another work [17], Gaussian mixture model (GMM), corresponding to each of six different manipulation types, is learned for image patches of small resolution ($16 \times 16$) and binary classification results are generated by comparing the average log-likelihood values of original and processed image patches. Although, the method is robust for tampering localization in uncompressed images, but the method does not address the problem in post-JPEG compressed images.

Recently, in [18], the residual-based local features are re-casted in form of convolutional neural network (CNN). Here, CNN is utilized to expose five different information-preserving manipulation types in image patches of resolution $128 \times 128$. Although, the method achieves good localization accuracies in uncompressed images, the problem of forgery localization in post-JPEG compressed images has not been addressed. In another recent work [19], CNN based multiple manipulation detector is proposed by designing filter bank to extract residual images to be input into CNN. The detector is proposed to detect 11 different manipulation types in uncompressed images of resolution $512 \times 512$. The seven convolutional layers based network pose limitations of large training database requirement and the detector is not explored to localize the forgeries in images.

Another general-purpose detector is proposed in [20] which target to learn features from interpolated images. The interpolated images are used as input into CNN with six convolutional layers and multi-class classification is performed to identify six different tampering operation types. The drawback of the method lies in the inability of the network to perform forgery localization in images. In another recent work [21], features are extracted from high-dimensional steganalysis based feature set
SRM (34, 671-dimensional) [24] and reduced to 714-dimensions for the purpose of general-purpose image forensics. The method is capable to detect 11 different manipulation types in uncompressed images. However, the performance is degraded for compressed images and for localization of forgeries in images. Thereafter, in [22], a constrained convolutional layer is proposed to highlight the specific manipulation features and then, CNN is employed to learn the features of five different manipulation types. The method achieves good binary and multi-class classification accuracies for uncompressed images but the performance degrades for compressed images. Furthermore, the method is not verified for forgery localization in images.

Recently, in [26], two CNNs are used in siamese CNN to simultaneously classify two image patches of size 150 × 150 as similarly processed or differently processed. The method addresses the global detection of five different manipulation types in images. In another work [27], the dense connectivity pattern between convolutional layers is used for better feature extraction to identify five different manipulation types in images. Further, in [28], global manipulations are identified using 8-convolutional layers based CNN which requires costly hardware to train the model for images of size 512 × 512. Thereafter, in [29], 11 different image manipulations are identified using a densely connected convolutional layers based CNN which is further utilized for forgery localization in images. Another CNN based general-purpose detector is proposed in [30] which utilizes depth-wise separable CNN to identify 11 different manipulation types.

Further, general-purpose image forensic method proposed in [31] extracts distinguishing features from local binary pattern (LBP) model and auto-regressive (AR) model employed with image residuals. The method is used to identify three different image smoothing operations. Recently, in [32], CNN is trained with reinforcement learning to identify 11 different image processing operations. Another recent method [33] claims that forensic detection accuracies are sensitive to the difference in pre-processing of training and testing images. Thus, a new method Gaussian mixture model resizing adaptation by fine-tuning (GRAFT) is proposed to identify five different image processing operations in small image patches of size 8 × 8. However, localization of forgeries in images under post-JPEG compression is not addressed.

As general-purpose image forensics is a challenging task, most of the existing methods take benefits either from rich and informative steganalysis based feature sets [16, 21, 31] or subtle tampering traces are extracted using CNN [18–20, 22, 26–30, 32]. Steganalysis based feature sets (SPAM [23], SRM [24] or LBP [25]) are developed based on noise residuals of images, and are enriched with information about image development process [18]. For this reason, good detection accuracy has been reported for various steganalysis based image forensic methods [5, 16, 21], but only for high resolution images. However, in [5, 16, 21], results have either not been reported for small image patches or have deteriorated significantly. Similarly, CNN based methods have extracted tampering artefacts using deep networks which facilitate better classification accuracy to detect multiple manipulations for large resolution images [19, 28]. Based on detailed study of existing general-purpose image forensic methods, the limitations of most of the existing works are: (1) high dimensionality of existing feature sets or computationally demanding deep CNN architectures, (2) rarely addressed forgery localization problem in post-JPEG compressed forged images and (3) rarely addressed fine-grain forgery localization (localization using small image patches) problem.

### 3 | PROPOSED APPROACH

#### 3.1 | Motivation

Neighbouring pixels of a natural image exhibit inherent statistical correlations [34], the correlations being function of the distance between the pixels in local region of image. These correlations are susceptible to the locally applied information-changing and information-preserving forgeries. Various image forensic methods for copy-move and splicing forgery localization have been developed by utilizing block feature matching techniques [35, 36], keypoint-based techniques [37, 38], local inconsistencies (illumination, geometrical and noise inconsistencies) based techniques [39, 40] and a combination of these techniques. These methods mainly focus to determine the statistical features of image blocks. Determined block features are matched with features of other blocks of the image to identify the inconsistent regions. Inspired by the success of forensic methods for the localization of information-changing forgeries in images, second-order joint statistics of image neighbouring blocks has been explored in the proposed work aimed at information-preserving forgery localization.

In the field of image forensics, most of the recent forensic methods [7, 21, 41] utilize residual domain image features rather than grey value domain [42] or difference domain [5, 6] image features. In previous works [7, 21], it has been demonstrated that content of images and JPEG compression artefacts are suppressed in residual domain as compared to grey value domain and difference domain images. Thus, residual domain images exhibit more clear artefacts of the forgery performed on the images. In the proposed work, the residual domain images are extracted by determining median filtered residual (MFR) [7] of an image. For a spatial domain image $I$, the MFR of image $I$ can be determined using

$$I_{res} = I - m_f\xi(I),$$

where $I_{res}$ denotes MFR of image $I$ and $m_f\xi(I)$ denotes median filtering operation, operates by replacing each pixel of image $I$ by the median of all pixel values in the neighbourhood region of size $\xi \times \xi$. In the proposed work, most commonly used window size $\xi = 3$ is selected to determine residual domain images. Subsequently, joint statistics of residual image blocks with neighbouring blocks is analyzed.

To analyze the efficacy of residual images as compared to grey value images, an example image is selected from BOWS2 [43]...
database and a locally forged image is created with forged region defined by $R_F \times C_F$ where $R_F = R$ and $C_F = \frac{C}{2}$, $R$ and $C$ being number of rows and columns in pristine example image. The forged region is created by Gaussian filtering the region $R_F \times C_F$ with window size $5 \times 5$ and standard deviation $\sigma = 1.1$ and then, forged image is JPEG compressed with $Q = 70$. The locally forged (also compressed) example image is shown in Figure 1(a). Further, MFR of the forged image (Figure 1(a)) is determined and shown in Figure 1(b). From Figure 1(b), a boundary between the forged and pristine regions of forged image MFR can be easily perceived while it is difficult to distinguish between the two regions in grey value domain forged image (Figure 1(a)).

Pixels in natural images exhibit certain statistical relationship with neighbouring pixels in each of the horizontal, vertical and diagonal directions. It is popular as 8-adjacency of a particular pixel $p$. Such relationship among pixels is destroyed in locally tampered images and consequently, the relative statistics of neighbouring blocks of an image is also disturbed. Thus, in order to extract information about relative statistics of reference block with adjacent blocks, 8-adjacent blocks are selected for each non-overlapping block in image MFR. For the blocks that exist at the boundary of image MFR, the neighbouring blocks are selected by padding the image MFR boundary with zeros matrix as shown in Figure 1(c). It should be noted that in Figure 1(c), $R_b$ and $C_b$ denote number of rows and columns in reference block (red block in Figure 1(c)) whereas $R$ and $C$ denote number of rows and columns in image MFR. Figure 1(d) demonstrates the selection of adjacent blocks (green blocks) for a particular non-overlapping reference block (red block).

Further, the features of each non-overlapping block (red block) are determined in terms of joint statistics of red block with neighbouring blocks. This feature set is used to obtain a trained model. Similarly, the features of each non-overlapping block of test image MFR (Figure 1(e)) are determined and each block is classified as original or forged using the trained model. The forged blocks are marked yellow in Figure 1(f). The complete flow diagram of proposed scheme is shown in Figure 2.

### 3.2 Joint statistics of residual image blocks and their neighbouring blocks

Considering each non-overlapping block of an image MFR and its 8-adjacent neighbours as a random variable, the joint statistics of the block under investigation and its neighbouring blocks is explored in terms of different association measures, that is cross-correlation and Euclidean distance between normalized histograms of two random variables. Denoting $R_b \times C_b$ sized reference block as $B$ and its $i^{th}$ neighbouring block of same size as $B^i$ in a locally forged image MFR, the cross-correlation between blocks $B$ and $B^i$ is defined as

$$\text{Cor}_{(B), (B^i)}(k, l) = \sum_{j=1}^{R_b} \sum_{z=1}^{C_b} R(j, z) B^i (y - k, z - l), \quad (2)$$

where

$$- (R_b - 1) \leq k \leq (R_b - 1),$$

$$- (C_b - 1) \leq l \leq (C_b - 1).$$
Here, \( i = 1, 2, \ldots, 8 \) correspond to 8-neighbour blocks to a reference block. The size of obtained cross-correlation matrix \( \text{Cor}(B_i \mid B^*) \) is \((2R_b - 1) \times (2C_b - 1)\). In order to extract maximum correlation between the block \( B \) and its \( i^{th} \) neighbouring block \( B^* \), maximum value is selected from the matrix \( \text{Cor}(B_i \mid B^*) \) using the equation

\[
\rho(B_i \mid B^*) = \max(\text{Cor}(B_i \mid B^*))
\] (3)

In addition to cross-correlation measure, Euclidean distance between normalized histograms of reference block \( B \) and its neighbouring blocks \( B^* \) (\( i = 1, 2, \ldots, 8 \)) is also used to analyze the statistical relationship between the neighbouring blocks of image MFR. Denoting \( h_B \) and \( h_{B^*} \) as normalized histograms of block \( B \) and its neighbouring block \( B^* \), respectively, the Euclidean distance between \( h_B \) and \( h_{B^*} \) is defined as

\[
D(B_i \mid B^*) = \sqrt[256]{\sum_{k=1}^{256} (h_B(k) - h_{B^*}(k))^2},
\] (4)

where \( i = 1, 2, \ldots, 8 \).

For 8-bit images, the histograms are plotted with 256 bins to analyze the grey values of reference block \( B \) with respect to grey values of each of its neighbouring blocks \( B^* \).

### 3.3 Comparison with existing approaches for forgery localization

In literature, the problem of general-purpose image forensics is mainly addressed using two approaches: (1) by utilizing rich steganalysis based features and (2) by employing CNN to extract high-level features corresponding to each manipulation type. Both approaches prove beneficial for global image forgery classification because the extracted features are enriched with detailed information about image development process. However, such detectors can be employed for forgery localization in images, considering each image block as an independent image, ignoring the fact that the image block is correlated with its neighbouring blocks. The features are extracted for independent small sized blocks, which suffer lack of available statistical information and result in significant performance drop for small sized image blocks. In contrast, the proposed method follows a significantly different approach and is specifically designed for forgery localization in images by extracting the features from joint statistics of each image block with eight of its neighbouring blocks. Furthermore, to the best of our knowledge, the proposed method is the first information-preserving forgery localization method that utilizes the inherent properties of neighbouring image blocks in terms of joint statistics of the blocks.

### 4 PROPOSED FEATURE SET \( F_{loc} \)

Based on the analysis given in preceding section, a new feature set is proposed for localization of various information-preserving manipulation types in images. The proposed feature set of 18-dimensions has following features:

#### 4.1 Maximum cross-correlation between reference block and its neighbouring blocks

The cross-correlation based proposed feature subset \( F_{\rho}^8 \) includes maximum cross-correlation values of block of image
MFR with each of its 8-neighbouring blocks and is defined as
\[
F^8_{Pr} = \rho_{(B_i|B^*)} \quad 1 \leq i \leq 8. \tag{5}
\]

Here, \(\rho_{(B_i|B^*)}\) denotes the maximum value of cross-correlation between reference block \(B\) and its \(i^\text{th}\) neighbouring block \(B^*\). Cross-correlation based feature subset \(F^8_{Pr}\) can be determined using (2) and (3).

### 4.2 Distance between normalized histograms of reference block and its neighbouring blocks

The similarity measure of locally forged reference blocks with their neighbouring blocks (forged or original) undergo substantial changes as compared to similarity between original blocks and their neighbouring blocks. To explore the similarities between two images, Jensen–Shannon (JS) divergence, Kullback-Leibler (KL) divergence and distance (city-block, Euclidean and Minkowski) between image normalized histograms are most commonly used measures. In the proposed work, the distance between normalized histograms of image MFR reference block \(B\) and its \(i^\text{th}\) neighbouring block \(B^*\) is determined using Euclidean distance. The Euclidean distance based 8-dimensional feature subset \(F^8_{D}\) is defined as
\[
F^8_{D} = D_{(B_i|B^*)} \quad 1 \leq i \leq 8. \tag{6}
\]

Here, \(D_{(B_i|B^*)}\) represents the distance between normalized histogram of image MFR reference block \(B\) and its neighbouring block \(B^*\). The distance is calculated using (4).

### 4.3 Statistical features of the reference block

In addition to the joint statistics of reference block and its adjacent blocks, the statistics of underlying reference block has also been explored in order to construct a robust feature set for forgery localization in images. Thus, in the proposed work, variance \(F_{V_B}\) and entropy \(F_{E_B}\) features of reference block \(B\) are included in the final feature set. The block variance and entropy based feature subsets \(F_{V_B}\) and \(F_{E_B}\) are defined as
\[
F_{V_B} = \frac{1}{(R_B \times C_B - 1)} \sum_{x=1}^{(R_B \times C_B)} (B_x - \mu_B)^2 \tag{7}
\]
\[
F_{E_B} = -b_B \log(b_B). \tag{8}
\]

Here, \(B_x, \mu_B\) and \(b_B\) denote \(x^{th}\) element of reference block \(B\), mean and normalized histogram of reference block \(B\) in image MFR, respectively.

### 4.4 Composite feature set

For the purpose of forgery localization in images, the feature subsets defined in Sections 4.1–4.3 are combined to form a 18-dimensional proposed feature set \(F_{bc}\) which can be defined as
\[
F_{bc} = [F^8_{Pr} \quad F^8_{D} \quad F_{V_B} \quad F_{E_B}]. \tag{9}
\]

## 5 DETAILS OF COMPARISON METHODS AND FEATURE SET ANALYSIS

The proposed method is compared against a steganalysis feature set based method SRM\(_{(T1,O3,q)}\) \cite{21} and a CNN based method MISLnet \cite{22}. The brief description of general-purpose information-preserving manipulation forensic methods SRM\(_{(T1,O3,q)}\) and MISLnet is as follows:

### 5.1 SRM\(_{(T1,O3,q)}\)

SRM\(_{(T1,O3,q)}\) \cite{21} feature set is designed using residual-based features rather than pixel-domain features to suppress the image contents and to enhance the forgery artefacts. The method takes advantage from one of the popular steganalysis based feature sets, spatial rich model (SRM) \cite{24}, by reducing its dimensions to 714 as compared to 34,671 dimensions of original SRM method. The feature set SRM is reduced to SRM\(_{(T1,O3,q)}\) by analyzing the average out of bag (OOB) error with respect to feature dimensionality. The reduced feature set is employed for global detection of 11 different manipulation types in different resolution uncompressed images. Results signify the significant performance drop of SRM\(_{(T1,O3,q)}\) for images of small resolution, that is \(32 \times 32\) and \(16 \times 16\). Additionally, the detection results using SRM\(_{(T1,O3,q)}\) are degraded for images under post-JPEG compression.

### 5.2 MISLnet

MISLnet \cite{22} is a convolutional neural network with a new constrained convolutional layer that facilitates suppression of image contents and enables CNN to learn distinguishing lower-level features for different manipulation types. Further, higher-level features are extracted from these lower-level features using a conventional deep CNN. Deeper convolutional layers in MISLnet use \(1 \times 1\) convolutions to determine additional associations between the feature maps determined from the previous layers. The method is employed for global detection of five different manipulation types in images. Results show that the average detection accuracy obtained using MISLnet is degraded under post-JPEG compression.

### 5.3 Feature sets efficacy comparison

In order to analyze the performance of proposed detector along with SRM\(_{(T1,O3,q)}\) \cite{21}, linear discriminant analysis (LDA) is performed. To this end, 1338 images are randomly selected from BOWS2 \cite{43} database and left vertical half of images \((R\times C_{F} \text{ sized region})\) are tampered using four different manipulation types: contrast enhancement with
FIGURE 3 LDA projections for post-JPEG compressed ($Q = 70$) 128 × 128 sized blocks manipulated with tampering groups (ori, $\alpha$, gau) and (ori, res, avg) using feature sets (a,b) $F_{loc}$ and (c,d) $SRM(T_{1,03,qp})$. The abbreviations $ori$, $\alpha$, $gau$, $res$ and $avg$ refer to original, contrast enhanced, Gaussian filtered, re-scaled and average filtered blocks.

Enhancement factor $\gamma = 0.5$, Gaussian filtering with standard deviation $\sigma = 1.1$ and window size $\xi = 5$, re-scaling with scaling factor ‘Scale=1.5’ and average filtering with window size $\xi = 5$. Then, the images are post-JPEG compressed with quality factor $Q = 70$. Further, feature sets $F_{loc}$ and $SRM(T_{1,03,qp})$ are determined for each of 128 × 128 sized non-overlapping blocks in the images. This results in five $F_{loc}$ and $SRM(T_{1,03,qp})$ feature sets each containing 10704 blocks corresponding to each of the four manipulation types and original image blocks. Further, the 2-D LDA projections of $F_{loc}$ and $SRM(T_{1,03,qp})$ features are determined for two manipulation groups: original (ori), contrast enhancement ($\alpha$), Gaussian filtering (gau) and [original (ori), re-scaling (res), average filtering (avg)]. The LDA projections for $F_{loc}$ and $SRM(T_{1,03,qp})$ are plotted in Figure 3.

Results shown in Figure 3 demonstrate the efficacy of proposed feature set $F_{loc}$ (Figure 3(a,b)) to classify between original image blocks and tampered image blocks as compared to $SRM(T_{1,03,qp})$ (Figure 3(c,d)) feature set, despite its smaller feature set dimensions. The same observation has already been made in one of the steganalysis based work [45] that only high dimensionality of employed feature set does not guarantee better performance. Here, LDA projections are determined only for four manipulation types, but similar behaviour can be observed for other manipulations too.

6 | EXPERIMENTAL METHODOLOGY

In this section, the database used for assessment of proposed detector $F_{loc}$ and methodology used to carry out experiments have been discussed.


### 6.1 Image datasets

The proposed feature set \( F_{30} \) is thoroughly evaluated to detect forged regions using images of four different databases: BOWS2 [43], BOSSBase [46], IEEE IFS-TC image forensics challenge [44] database and a public database [47]. The usage of these databases in the proposed work is inspired by state-of-the-art methods [17, 21, 22]. BOWS2 and BOSSBase databases contain 10,000 .pgm format images of resolution 512 × 512, containing vast variety of natural scenes and low frequency images. These databases have been used in ‘Break our water-marking system’ contest for the first time. In the proposed work, we have used 2000 images from each of BOWS2 and BOSSBase databases. IEEE IFS-TC database contains high resolution pristine images, mostly having size 1024 × 768 or 768 × 1024, along with copy-move forged or spliced images and their binary masks. We have used 1013 pristine images from the database and centrally cropped the images to size 512 × 512. The public database contains wide variety of uncompressed images having size 4272 × 2848. In the proposed work, 25 images from public database are used to centrally crop 40 tiles of size 512 × 512 from each image. Out of these 1000 image tiles of size 512 × 512, 987 tiles are selected and merged with the 1013 images of IEEE IFS-TC database to create a dataset (IEEE+HR) which contains 2000 images of size 512 × 512. Further, following different locally tampered image datasets are created for DB ∈ {BOWS2, BOSSBase, (IEEE+HR)} to determine the forgery localization results:

(i) \( CB(BOWS2)^{\delta}_{\text{ori+Q}} \): To create \( CB(BOWS2)^{\delta}_{\text{ori+Q}} \) dataset containing completely original blocks/forged blocks \( CB \), 1338 images are randomly selected from BOWS2 database and left vertical half region of each image is tampered with the manipulation types \( \delta \) listed in Table 1 for both the parameters shown. It should be noted that half of the image is being tampered so that a balanced dataset, with equal number of original and forged image blocks, can be obtained for proper training of the classifier. Resulting images are post-JPEG compressed with quality factor \( Q = 70, 30 \). The locally tampered images, thus created, are divided into non-overlapping blocks of size \( \delta \in \{128 \times 128, 64 \times 64, 32 \times 32 \} \) to create \( CB(BOWS2)^{\delta}_{\text{ori+Q}} \). Each block of dataset \( CB(BOWS2)^{\delta}_{\text{ori+Q}} \) is labelled as original (ori) or forged. It should be noted that all the blocks in dataset \( CB(BOWS2)^{\delta}_{\text{ori+Q}} \) extracted from images of resolution 512 × 512, are either completely original or completely forged. Training set for binary classification is defined as \( \{ CB(BOWS2)^{\delta}_{\text{ori+Q}} \} \cup \{ CB(BOWS2)^{\delta}_{\text{ori+Q}} \} \) whereas testing results are determined for complementary images in the set.

To analyze the effect of different local forged regions in images, two variants of \( CB(BOWS2)^{\delta}_{\text{ori+Q}} \) are created. One of the variant is created by tampering selected 1338 images for image rows 65–448, columns 5–384 and images are post-JPEG compressed with \( Q = 30 \). Another variant is created by tampering 669 images globally along with 669 original images and all 1338 images are post-JPEG compressed with \( Q = 30 \). It should be noted that tampering operations \( \delta \) are taken from Table 1 with bold marked parameters. Further, non-overlapping blocks of size 64 × 64 are cropped from both the variants to create two different datasets \( CB(BOWS2\text{var})^{\delta}_{\text{ori+Q}} \) and \( CB(BOWS2\text{var}2)^{\delta}_{\text{ori+Q}} \) respectively.

(ii) \( CB(DB)^{\delta}_{\text{ori+Q}} \): In order to analyze the robustness of proposed approach with respect to different databases, \( CB(DB)^{\delta}_{\text{ori+Q}} \) is created for images of DB database where DB ∈ \{BOSSBase, (IEEE+HR)\}. The 1338 images are randomly selected from DB \( DB \in \{ \text{BOSSBase, (IEEE+HR)} \} \) and left vertical half of images are tampered with manipulations \( \delta \) listed in Table 1 using the bold marked parameters. Further, the images are post-JPEG compressed with \( Q = 30 \). Then, images are divided into non-overlapping blocks of size 64 × 64 to create dataset \( CB(DB)^{\delta}_{\text{ori+Q}} \) for DB ∈ \{BOSSBase, (IEEE+HR)\}.

(iii) \( CB(PB)^{\delta}_{\text{ori+Q}} \): In practical scenarios, a forensic analyzer usually encounters locally forged images with partially forged blocks. Thus, in order to obtain precise forgery localization detector, training dataset is created by including partially forged blocks/original blocks along with completely forged blocks/original blocks. Partially forged blocks/original blocks dataset \( CB(PB)^{\delta}_{\text{ori+Q}} \) is created for DB ∈ \{BOSSBase, (IEEE+HR)\} dataset. Among 2000 images of dataset DB, 1000 images are locally tampered in left vertical half region of images, 500 images are locally tampered for image rows, columns 97–448 and rest 500 images are locally tampered for image rows, columns 81–448. The images are tampered with manipulations \( \delta \) given in Table 1 using bold marked parameters and post-JPEG compressed with quality factor \( Q = 70, 30 \). Further, images are divided into non-overlapping blocks of size 64 × 64 which constitute the set \( CB(PB)^{\delta}_{\text{ori+Q}} \) for DB ∈ \{BOSSBase, (IEEE+HR)\}. Along with completely forged blocks/original blocks, the constructed
Experimental setup

The proposed detector $F_{\text{loc}}$ is evaluated by developing the trained models using the datasets $CB\{DB\}^{S|\text{op}+Q}$ and $PB\{DB\}^{S|\text{op}+Q}$, where $DB\in\{\text{BOWS2, BOSSBase, (IEEE+HR)}\}$. For dataset $CB\{DB\}^{S|\text{op}+Q}$ results are presented in terms of 70% − 30% holdout validation accuracies (i.e. 70% images are used for training and 30% are used for testing), achieved using ensemble of decision tree classifiers with bagging. To achieve a generalized detector, the results are determined for $21,408, 85,632$ and $342,528$ image blocks for $S = 128 \times 128, 64 \times 64, 32 \times 32$, respectively. The proposed detector $F_{\text{loc}}$ results are compared with two state-of-the-art detectors SRM$_{(T|O,3+Q)}$ [21] and MISLnet [22]. Detector SRM$_{(T|O,3+Q)}$ is a 714-dimensional statistical features based multiple manipulation detector whereas MISLnet is a CNN-based multiple manipulation detector.

For compilation of SRM$_{(T|O,3+Q)}$ results, the above defined dataset $CB\{\text{BOWS2}\}^{S\|\text{op}+Q}$ for $S \in \{128 \times 128, 64 \times 64, 32 \times 32\}$ is used to determine 70% − 30% holdout validation results using ensemble classifier [48]. On the other hand, MISLnet, being a CNN based detector, require large number of images to optimize the network parameters. Thus, following the experimental methodology and network settings described in [22], MISLnet is trained for dataset $CB\{\text{BOWS2}\}^{S\|\text{op}+Q}$ images for $S = 64 \times 64$ with 85632 image blocks. However, for block size $S = 128 \times 128$, number of available blocks (21408) are not sufficient to train MISLnet. Therefore, 1013 large resolution (1024 × 768, 1024 × 683 and 1024 × 585) images are randomly selected from IEEE IFS-TC image forensics challenge [44] database and left vertical half region of each image is tampered with the manipulations (op) listed in Table 1 for both the parameters given. Resulting images are post-JPEG compressed.
TABLE 3  Classification results for compressed (Q) forged blocks of resolution S against compressed original blocks of same resolution. Proposed detector results, in terms of classification accuracy $C_{acc}$ (%) are compared with SRM$_{(T_1,0,3,gt)}$ [21] and MISLnet [22] results. Best results are highlighted with grey colour.

| $S \times Q$ | Method/Classifier | AVG $(3 \times 3)$ | AVG $(5 \times 5)$ | RES $(1.2)$ | RES $(1.5)$ | AWGN $(0.0005)$ | AWGN $(0.001)$ | UM $(0.2)$ | UM $(0.4)$ | JPP $(1.5)$ | JPP $(2.0)$ | Average |
|---------------|-------------------|------------------|------------------|-------------|-------------|-----------------|-----------------|-------------|-------------|-------------|-------------|-----------|
| $128 \times 128$ | $F_{ao}$ | 99.64 | 99.78 | 92.48 | 95.38 | 94.27 | 97.34 | 83.74 | 87.87 | 83.17 | 82.54 | 91.61 |
| | SRM$_{(T_1,0,3,gt)}$ | 99.49 | 99.74 | 86.03 | 91.51 | 96.64 | 99.60 | 61.96 | 70.52 | 52.27 | 67.99 | 82.58 |
| | MISLnet/ERT | 97.07 | 98.53 | 83.98 | 86.02 | 95.89 | 98.65 | 60.81 | 66.29 | 58.32 | 57.30 | 80.29 |
| | MISLnet/Softmax | 97.10 | 98.35 | 83.88 | 86.26 | 95.75 | 98.60 | 61.31 | 66.45 | 58.13 | 57.46 | 80.33 |
| $30$ | $F_{ao}$ | 99.16 | 99.53 | 91.09 | 94.41 | 87.36 | 91.42 | 84.55 | 87.39 | 82.79 | 83.10 | 90.08 |
| | SRM$_{(T_1,0,3,gt)}$ | 98.85 | 99.49 | 81.49 | 88.52 | 88.45 | 94.69 | 60.42 | 67.14 | 51.46 | 66.23 | 79.67 |
| | MISLnet/ERT | 93.78 | 97.21 | 79.81 | 91.46 | 88.87 | 95.10 | 58.35 | 66.35 | 57.56 | 55.21 | 78.40 |
| | MISLnet/Softmax | 93.71 | 97.27 | 79.68 | 91.42 | 88.76 | 95.18 | 58.03 | 66.18 | 57.17 | 55.61 | 78.30 |
| $64 \times 64$ | $F_{ao}$ | 93.64 | 96.92 | 78.81 | 85.27 | 91.67 | 95.38 | 68.32 | 71.64 | 65.90 | 66.29 | 81.40 |
| | SRM$_{(T_1,0,3,gt)}$ | 93.32 | 95.43 | 81.58 | 87.41 | 93.49 | 98.17 | 59.14 | 65.97 | 51.99 | 51.14 | 77.76 |
| | MISLnet/ERT | 95.54 | 97.09 | 84.46 | 90.58 | 89.39 | 95.96 | 59.33 | 65.12 | 53.23 | 53.92 | 78.46 |
| | MISLnet/Softmax | 95.49 | 97.17 | 84.60 | 90.44 | 89.43 | 95.69 | 58.55 | 64.44 | 53.31 | 54.35 | 78.35 |
| $30$ | $F_{ao}$ | 92.05 | 96.41 | 76.58 | 83.53 | 81.36 | 87.33 | 68.68 | 72.04 | 65.77 | 66.20 | 79.00 |
| | SRM$_{(T_1,0,3,gt)}$ | 90.90 | 93.69 | 75.91 | 82.50 | 82.81 | 90.20 | 58.35 | 64.27 | 51.11 | 51.02 | 74.08 |
| | MISLnet/ERT | 93.13 | 95.64 | 79.88 | 87.07 | 77.87 | 84.82 | 58.40 | 65.51 | 53.25 | 53.79 | 74.94 |
| | MISLnet/Softmax | 93.42 | 95.66 | 80.26 | 87.27 | 77.77 | 84.43 | 58.44 | 64.96 | 52.85 | 53.83 | 74.89 |
| $32 \times 32$ | $F_{ao}$ | 87.58 | 93.10 | 71.56 | 79.18 | 89.24 | 93.89 | 61.00 | 64.70 | 58.31 | 58.23 | 75.68 |
| | SRM$_{(T_1,0,3,gt)}$ | 86.95 | 89.83 | 75.35 | 82.65 | 88.36 | 95.10 | 56.70 | 62.10 | 51.16 | 51.23 | 73.94 |
| | MISLnet/Softmax | 84.85 | 91.61 | 68.86 | 76.06 | 78.48 | 84.68 | 60.23 | 64.29 | 57.50 | 57.65 | 72.42 |
| | SRM$_{(T_1,0,3,gt)}$ | 82.87 | 86.96 | 69.34 | 76.23 | 74.99 | 83.82 | 56.43 | 61.07 | 50.64 | 50.60 | 69.30 |

The results are determined for image block size $64 \times 64$ and compression quality factor $Q = 30$.

TABLE 4  Comparison of localization results obtained using $F_{ao}$ for different databases

| Dataset | Average accuracy (%) | Average F1 score |
|---------|----------------------|------------------|
| CB$(BOWS2)^{64}_{(eff)}+Q^{30}$ | 81.76 | 0.8153 |
| CB$(BOSSBase)^{64}_{(eff)}+Q^{30}$ | 78.48 | 0.7826 |
| CB$(IEEE+HR)^{64}_{(eff)}+Q^{30}$ | 79.04 | 0.7871 |

The results are determined for image block size $64 \times 64$ and compression quality factor $Q = 30$.

TABLE 5  Comparison of localization results obtained using $F_{ao}$ for different forgery locations in images

| Dataset | Average accuracy (%) | Average F1 score |
|---------|----------------------|------------------|
| CB$(BOWS2)^{64}_{(eff)}+Q^{30}$ | 81.76 | 0.8153 |
| CB$(BOWS2var1)^{64}_{(eff)}+Q^{30}$ | 75.92 | 0.7555 |
| CB$(BOWS2var2)^{64}_{(eff)}+Q^{30}$ | 82.30 | 0.8229 |

The results are determined for image block size $64 \times 64$ and compression quality factor $Q = 30$.

21408 blocks of $CB$(BOWS2)$^{128}_{(eff)}+Q^{30}$ dataset to create training set $\{CB$(BOWS2$+IEEE)^{128}_{(eff)}+Q^{30}$\}$.

The MISLnet results are also presented in terms of $70\% - 30\%$ holdout validation accuracies achieved. Additionally, MISLnet network has the limitation that it is not designed to determine results for image resolution $32 \times 32$ as network parameters need to be modified for this purpose. Thus, proposed detector $F_{ao}$ results are not compared with MISLnet results for image block size $32 \times 32$.

Further, in order to obtain more precise localization performance, proposed detector $F_{ao}$, along with SRM$_{(T_1,0,3,gt)}$ and MISLnet, is trained using partially forged blocks dataset $DB^{128}_{(eff)}$ with 128,000 image blocks where $DB \in \{BOWS2, BOSSBase, (IEEE+HR)\}$ and $Q = 70, 30$. The results are reported in terms of $70\% - 30\%$ holdout validation accuracies and F1 scores obtained by all the three detectors for image block size $64 \times 64$.

The localization results for proposed detector $F_{ao}$ are reported using MATLAB 2018(a) classification learner application. MISLnet results are compiled using Caffe deep learning framework with python 3.6. Number of training iterations are set to 47,000 and results are determined using two classifiers: softmax and extremely randomized trees (ERT) as mentioned in original manuscript [22].
TABLE 6  Localization results of partially forged/partially original blocks of resolution 64 x 64 in post-JPEG compressed (Q = 30) images of different datasets. Proposed detector results are compared with SRM(T1,O3,q∗) [21] and MISLnet [22] results. Best results are highlighted with grey colour

| Dataset               | Method/Classifier | Parameters | MF (5x5) | CE (0.5) | GAU (5x5) | WNR (5x5) | JPEG (Q=90) | AVG (5x5) | RES (1.5) | AWGN (0.0001) | UM (0.4) | JP2 (1.5) | Average |
|-----------------------|-------------------|------------|----------|----------|-----------|-----------|-------------|-----------|-----------|----------------|-----------|-----------|---------|
| ![BOSS2](64)          | F_∞               | C_acc(%)   | 84.66    | 70.54    | 82.37     | 77.25     | 66.13       | 90.85     | 78.04     | 86.00          | 68.66    | 65.39     | 76.99   |
|                       |                   | F1 score   | 0.8622   | 0.7436   | 0.8427    | 0.8002    | 0.7093      | 0.9150    | 0.8053    | 0.8740         | 0.7175   | 0.7670    | 0.8037  |
| SRM(T1,O3,q∗)         | C_acc(%)          | 82.54      | 63.54    | 83.45     | 84.16     | 59.95      | 93.40       | 82.17     | 89.30     | 76.08          | 67.08    | 54.76     | 76.04   |
|                       | F1 score          | 0.8238     | 0.6245   | 0.8293    | 0.8391    | 0.5794     | 0.9323      | 0.8391    | 0.8954    | 0.6235         | 0.5555   | 0.7542    | 0.7452  |
| MISLnet/ ERT          | C_acc(%)          | 84.33      | 64.60    | 86.50     | 83.39     | 54.80      | 94.51       | 83.79     | 0.9192    | 64.75          | 53.76    | 76.24     | 76.24   |
|                       | F1 score          | 0.8298     | 0.5950   | 0.8573    | 0.8185    | 0.4907     | 0.9422      | 0.8200    | 0.919     | 0.6377         | 0.4558   | 0.7359    | 0.7359  |
| MISLnet/ softmax      | C_acc(%)          | 84.28      | 64.11    | 84.38     | 82.88     | 54.95      | 94.77       | 83.46     | 91.95     | 65.01          | 54.52    | 76.03     | 76.03   |
|                       | F1 score          | 0.8302     | 0.5936   | 0.8286    | 0.8255    | 0.4967     | 0.9451      | 0.8140    | 0.9105    | 0.6573         | 0.4399   | 0.7341    | 0.7341  |
| ![BOSSBase](64)       | F_∞               | C_acc(%)   | 80.63    | 68.33    | 79.00     | 73.14     | 64.11       | 83.22     | 73.46     | 86.73          | 66.00    | 63.85     | 73.56   |
|                       |                   | F1 score   | 0.8198   | 0.7120   | 0.8136    | 0.7699    | 0.6943      | 0.8487    | 0.7680    | 0.8784         | 0.7680   | 0.6923    | 0.7765  |
| SRM(T1,O3,q∗)         | C_acc(%)          | 78.11      | 61.29    | 79.84     | 79.10     | 57.89      | 88.47       | 76.35     | 89.02     | 61.64          | 53.54    | 72.53     | 72.53   |
|                       | F1 score          | 0.7831     | 0.5923   | 0.8006    | 0.7912    | 0.5614     | 0.8828      | 0.7704    | 0.8930    | 0.5335         | 0.5411   | 0.7169    | 0.7169  |
| MISLnet/ ERT          | C_acc(%)          | 80.38      | 63.07    | 80.95     | 79.80     | 54.43      | 91.19       | 78.43     | 91.19     | 60.73          | 53.45    | 73.36     | 73.36   |
|                       | F1 score          | 0.7831     | 0.6138   | 0.7870    | 0.7802    | 0.4712     | 0.9071      | 0.7580    | 0.9048    | 0.6030         | 0.4549   | 0.7063    | 0.7063  |
| MISLnet/ softmax      | C_acc(%)          | 84.05      | 68.33    | 81.19     | 79.83     | 55.46      | 91.25       | 78.50     | 90.03     | 60.76          | 53.94    | 73.47     | 73.47   |
|                       | F1 score          | 0.7970     | 0.6138   | 0.7874    | 0.7811    | 0.4491     | 0.9079      | 0.7568    | 0.8970    | 0.6172         | 0.4459   | 0.7035    | 0.7035  |
| ![IEEE+HR](64)        | F_∞               | C_acc(%)   | 81.83    | 69.49    | 81.54     | 76.55     | 67.41       | 86.80     | 76.64     | 86.84          | 68.66    | 66.96     | 76.27   |
|                       |                   | F1 score   | 0.8394   | 0.7359   | 0.8361    | 0.7943    | 0.7173      | 0.8794    | 0.7939    | 0.8811         | 0.7274   | 0.7164    | 0.7921  |
| SRM(T1,O3,q∗)         | C_acc(%)          | 82.24      | 60.18    | 82.98     | 83.60     | 56.84      | 92.35       | 74.96     | 88.88     | 59.35          | 51.35    | 73.27     | 73.27   |
|                       | F1 score          | 0.8308     | 0.5639   | 0.8361    | 0.8357    | 0.5367     | 0.9226      | 0.7656    | 0.8917    | 0.5396         | 0.4753   | 0.7198    | 0.7198  |
| MISLnet/ ERT          | C_acc(%)          | 84.07      | 60.95    | 81.79     | 86.16     | 54.62      | 94.14       | 75.34     | 91.85     | 55.68          | 51.63    | 73.64     | 73.64   |
|                       | F1 score          | 0.8294     | 0.5641   | 0.7802    | 0.8466    | 0.4810     | 0.9381      | 0.6988    | 0.9117    | 0.5485         | 0.3822   | 0.6978    | 0.6978  |
| MISLnet/ softmax      | C_acc(%)          | 84.27      | 60.74    | 81.93     | 86.37     | 55.21      | 94.12       | 75.93     | 91.96     | 56.31          | 52.62    | 73.95     | 73.95   |
|                       | F1 score          | 0.8194     | 0.5475   | 0.7786    | 0.8460    | 0.4729     | 0.9381      | 0.6988    | 0.9102    | 0.5374         | 0.2692   | 0.6818    | 0.6818  |
7 | RESULTS AND DISCUSSION

The forgery localization results in terms of classification accuracies achieved by proposed detector $F_{loc}$, SRM$_{(T1,O3,qb)}$ and MISLnet are reported using the formula

\[ C_{acc} = \frac{TP + TN}{P + N}. \]  

(10)

Here, $TP$ and $TN$ denote number of true positive and number of true negative, respectively. Also, $P$ and $N$ denote total number of positive and negative samples, respectively. Additionally, the F1-score of classification is also determined using the formula

\[ \text{F1 score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \]  

(11)

where Precision and Recall are defined as

\[ \text{Precision} = \frac{TP}{TP + FP}, \]  

(12)

\[ \text{Recall} = \frac{TP}{TP + FN}. \]  

(13)

Here, $FP$ and $FN$ denote number of false positive and number of false negative, respectively.

7.1 | Forgery detection in completely forged/completely original blocks

The localization results for images containing either completely forged blocks or completely original blocks are simulated for $CB$($\text{BOWS2}$)${^64}_{qsp/\text{ori}+Q}$ dataset images using proposed detector $F_{loc}$ for $S \in \{128 \times 128, 64 \times 64, 32 \times 32\}$. The proposed detector results are compared with the classification accuracies achieved by existing detectors SRM$_{(T1,O3,qb)}$ [21] and MISLnet [22]. The binary classification results are reported for 20 different test case scenarios, including 10 different manipulation types with two parameters each, parameters being listed in Table 1. The results are tabulated in Tables 2 and 3. Results demonstrate that the proposed detector performs better than existing detectors in most of the test case scenarios, specifically for the images compressed with low quality factor of compression ($Q = 30$). The proposed detector $F_{loc}$ exhibits significant performance improvement for detection of contrast enhancement, unsharp masking, JPEG compression and JPEG2000 compression as compared to state-of-the-art detectors. Additionally, the proposed detector $F_{loc}$ achieves maximum average classification accuracy to localize forged regions in compressed images. It should be noted that MISLnet results are simulated for block size $S \in \{128 \times 128, 64 \times 64\}$ as for $S = 32 \times 32$, MISLnet network needs modification. As discussed in the previous section, MISLnet results for $S = 128 \times 128$ are simulated for images of dataset $CB$($\text{BOWS2+IEEE}$)$^{128}_{qsp/\text{ori}+Q}$ for 67,810 image blocks to properly train the network.

7.2 | Robustness of $F_{loc}$ results for different databases and different forged region locations

We have analyzed the robustness of proposed detector $F_{loc}$ results with respect to different databases. To this end, the results are determined for dataset $CB$($DB$)${^{64}}_{qsp/\text{ori}+Q}$ (creation of $CB$($DB$)${^{64}}_{qsp/\text{ori}+Q30}$ is discussed in Section 6.1) where $DB \in \{\text{BOWS2}, \text{BOSSBase}, (\text{IEEE+HR})\}$. The 70 – 30% holdout validation results are determined for manipulation types ($qP$) listed in Table 1 for bold marked parameters. The average accuracies and average F1 scores for databases BOWS2, BOSSBase and (IEEE+HR) are shown in Table 4. It can be seen from the table that the results obtained for different databases do not vary much.

Further, the robustness of the proposed detector with respect to different forged region locations is analyzed. This is achieved by determining and comparing the results for previously defined datasets $CB$($BOWS2$)${^{64}}_{qsp/\text{ori}+Q30}$, $CB$($BOWS2var1$)${^{64}}_{qsp/\text{ori}+Q30}$ and $CB$($BOWS2var2$)${^{64}}_{qsp/\text{ori}+Q30}$ (datasets are discussed in Section 6.1). The 70% – 30% holdout validation results are determined for three datasets for the manipulations ($qP$) listed in Table 1 using bold marked parameters. The average accuracies and average F1 scores for three cases are listed in Table 5 which show that the results are degraded for the dataset $CB$($BOWS2var1$)${^{64}}_{qsp/\text{ori}+Q30}$ when blocks of 669 globally forged images are considered as forged and blocks of other 669 images are considered as original. The reason for performance drop in this case is that training dataset contains examples of forged blocks surrounded by only forged neighbouring blocks and original blocks surrounded by only original neighbouring blocks. For the remaining two cases, training datasets are generalized datasets containing forged blocks with original and forged neighbouring blocks and vice-versa. Thus, we can conclude that dataset created by including forged image blocks and original image blocks from different images is not suitable for forgery localization using $F_{loc}$, rather it is the most common method to create dataset for global forgery detection in images.

7.3 | Forgery detection in partially forged/partially original blocks

To analyze the performance of proposed detector for its applicability in real world scenarios, the localization results are determined for compressed images containing partially forged or partially original blocks. In order to achieve this, an ensemble of 30 decision trees has been trained with images of datasets $PB$($DB$)${^{64}}_{qsp/\text{ori}+Q30}$ for $DB \in \{\text{BOWS2}, \text{BOSSBase}, (\text{IEEE+HR})\}$, containing partially forged and partially original blocks. It should be noted that image blocks with half or more forged region is labelled as forged otherwise labelled as original. Trained models using proposed detector $F_{loc}$ state-of-the-art detectors SRM$_{(T1,O3,qb)}$ [21] and MISLnet [22] are developed for the operations ($qP$) listed in Table 1 with bold marked parameters.
Figure 4 (a) Locally forged image (information-changing forgery) from database [44], (b) ground-truth of the image, (c) locally forged compressed ($Q = 70$) image with red marked average filtered ($\xi = 5$) blocks, forgery localization results using (d) proposed detector $F_{loc}$, (e) SRM($T_1, O_3, q^*$) [21] and (f) MISLnet [22]

Figure 5 (a) Locally forged image (information-changing forgery) from database [44], (b) ground-truth of the image, (c) locally forged compressed ($Q = 70$) image with red marked median filtered ($\xi = 5$) blocks, forgery localization results using (d) proposed detector $F_{loc}$, (e) SRM($T_1, O_3, q^*$) [21] and (f) MISLnet [22]

The 70 – 30% holdout validation results are determined for proposed feature set $F_{loc}$ along with SRM($T_1, O_3, q^*$) and MISLnet. The results are compiled in Table 6 in terms of the detection accuracies $C_{acc}$ and F1 scores achieved by different methods. Results in Table 6 show that the proposed detector achieves better results as compared to state-of-the-art detectors, especially for the detection of median filtering, contrast enhancement, unsharp masking, JPEG and JPEG2000 compression. Additionally, the proposed detector achieves best average accuracies and average F1-scores for the three databases.

To serve one’s purposes, the malicious user usually create locally tampered images by modifying the content of the image and by performing post-processing operations to hide the footprints of underlying forgery. IEEE IFS-TC image forensics challenge [44] database contains images created using copy-move or splicing forgery. Two images have been selected from the database and shown in Figures 4(a) and 5(a) along with respective ground-truths in Figures 4(b) and 5(b). These images have been created using copy-move or splicing forgery. Further, the images shown in Figures 4(a) and 5(a) are locally average filtered and median filtered with window size $\xi = 5$, respectively and complete images are post-JPEG compressed with quality factor $Q = 70$. The regions for locally applied average filtering and median filtering have been selected such that it correspond
Figure 6 (a) Original image 1 from BOWS2 [43] database, (b) original image 2 from BOWS2 [43] database, (c) locally forged compressed ($Q = 70$) image with red marked Gaussian filtered ($\xi = 5, \sigma = 1.1$) blocks, forgery localization results using (d) proposed detector $F_{loc}$, (e) SRM($T_1, O_3, q^*$) [21] and (f) MISLnet [22] to minimum covering area for the primary forgery. The average filtered and median filtered blocks are marked with red colour and are shown in Figures 4(c) and 5(c), respectively. It is worth mentioning here that blocks with half or more forged pixels are considered as forged otherwise considered as original. Then, trained models using feature sets $F_{loc}$, SRM($T_1, O_3, q^*$) and MISLnet are used to identify locally forged blocks of size $64 \times 64$ in test images given in Figures 4(c) and 5(c) and results are shown in Figures 4(d–f) and 5(d–f), respectively. Along with the visual results, the detection accuracies ($C_{acc}$) and F1 scores for Figures 4(d–f) and 5(d–f) are determined. Results ascertain that the proposed detector $F_{loc}$ outperforms state-of-the-art detectors SRM($T_1, O_3, q^*$) [21] and MISLnet [22] to localize average filtered and median filtered blocks of size $64 \times 64$ in compressed ($Q = 70$) images despite its smaller feature set dimensions (18-dimensions).

Additionally, to ensure localization of Gaussian filtering and re-scaling in images, two test images are created with local Gaussian filtered and re-scaled regions. To this end, four images are selected from BOWS2 [43] database as shown in Figures 6(a,b) and 7(a,b). The image in Figure 6(b) is Gaussian filtered with window size $\xi = 5$, standard deviation $\sigma = 1.1$ and image in Figure 7(b) is rescaled with a factor of 1.5 followed by cropping to same size. Then, a portion of manipulated versions of Figures 6(b) and 7(b) are copied and pasted onto Figures 6(a) and 7(a), respectively at same locations. The resulting images...
are post-JPEG compressed with quality factor $Q = 70$ and are shown in Figures 6(c) and 7(c). The red marked blocks in Figures 6(c) and 7(c) denote Gaussian filtered and re-scaled blocks, respectively. The test images, created in this way, are tested using detectors $F_{loc}$, SRM$(T_{1},O_{3},q^{*})$, MISLnet and results are shown in Figures 6(d–f), 7(d–f) for localization of Gaussian filtered blocks and re-scaled blocks, respectively. Results, in terms of detection accuracies and F1 scores, exhibit better performance of proposed detector $F_{loc}$ with low feature set dimensions (18-dimensions) as compared to existing detectors SRM$(T_{1},O_{3},q^{*})$ [21] and MISLnet [22].

### 7.4 Results analysis

Based on the results determined in Sections 7.1–7.3, following observations can be made about the efficacy of proposed localization detector $F_{loc}$ as compared to state-of-the-art detectors:

1. For forgery localization in images containing completely forged (or completely original) blocks, the main contribution of proposed detector $F_{loc}$ lies in its better performance for block size $128 \times 128$ compressed with two quality factors of compression $Q$ ($Q = 30, 70$). For smaller block sizes $64 \times 64$ and $32 \times 32$, performance of detector $F_{loc}$ degrades for specific manipulation types (Tables 2 and 3). This is due to decrease in available information with decrease in block size. However, average accuracies achieved by detector $F_{loc}$ are higher as compared to SRM$(T_{1},O_{3},q^{*})$ [21] and MISLnet [22] detectors for all block resolutions ($S \in \{128 \times 128, 64 \times 64, 32 \times 32\}$) and for both values of compression quality factor $Q$ ($Q = 70, 30$). Additionally, the proposed detector $F_{loc}$ is robust to images from various databases and from different locations of forged regions in images.

2. For forgery localization in images containing partially forged (or partially original) blocks, the proposed detector $F_{loc}$ exhibits comparable or better performance as compared to existing detectors SRM$(T_{1},O_{3},q^{*})$ and MISLnet. The holdout validation (Table 6) and testing results (Figures 4–7) show that the detector $F_{loc}$ is suitable for fine-grain forgery localization in compressed images despite its small dimensions (18-dimensions).

### 8 CONCLUSION

Determination of precise location of forgery in images is very crucial from forensic point of view which set-up new research goals to localize information-changing and information-preserving manipulations in images. In the past years, a lot of work has been done for localization of information-changing manipulations [3, 4] whereas localization of information-preserving manipulations [14, 15, 17] is still an emerging area of research. This paper presents a new method to localize 10 different information-preserving manipulation types in original, spliced and copy-move forged images of different databases in compressed format. A 18-dimensional new feature set $F_{loc}$ is proposed, by analyzing the joint statistics of image non-overlapping blocks with 8-neighbouring blocks, to localize average filtering, median filtering, Gaussian filtering, Wiener filtering, JPEG compression, JPEG2000 compression, unsharp masking, contrast enhancement, additive Gaussian noise addition and re-scaling operations. The proposed detector $F_{loc}$ results are compared with two recent multiple manipulation detectors SRM$(T_{1},O_{3},q^{*})$ [21] and MISLnet [22]. Results demonstrate the effectiveness of detector $F_{loc}$ despite its lower feature set dimensions as compared to high feature set dimensions of SRM$(T_{1},O_{3},q^{*})$ (714-dimensions) and complex CNN network of MISLnet (5-layers).

#### REFERENCES

1. Kirchner, M., Bohme, R.: Hiding traces of resampling in digital images. Trans. Info. For Sec. 3(4), 582–592 (2008)
2. Stamm, M.C., Liu, K.J.R.: Anti-forensics of digital image compression. Trans. Info. For Sec. 6(3), 1050–1065 (2011)
3. Cozzolino, D., Poggi, G., Verdoliva, L.: Copy-move forgery detection based on patchmatch. In: 2014 IEEE International Conference on Image Processing (ICIP), Paris (2014)
4. Zhao, X., et al.: Passive image-splicing detection by a 2D noncausal Markov model. IEEE Trans. Circuits Syst. Video Technol. 25(2), 185–199 (2015)
5. Kirchner, M., Fridrich, J.: On detection of median filtering in digital images. Proc. SPIE 7541, 754110, (2010)
6. Chen, C., Ni, J., Huang, J.: Blind detection of median filtering in digital images: A difference domain based approach. IEEE Trans. Image Process. 22(12), 4699–4710 (2013)
7. Kang, X., et al.: Robust median filtering forensics using an autoregressive model. IEEE Trans. Inf. Forensics Secur. 8(9), 1456–1468 (2013)
8. T. H., et al.: JPG quantization step estimation and its applications to digital image forensics. IEEE Trans. Inf. Forensics Secur. 12(1), 123–133 (2017)
9. Stamm, M.C., Liu, K.J.R.: Forensic detection of image manipulation using statistical intrinsic fingerprints. IEEE Trans. Inf. Forensics Secur. 5(3), 492–506 (2010)
10. Cao, G., et al.: Contrast enhancement-based forensics in digital images. IEEE Trans. Inf. Forensics Secur. 9(3), 515–525 (2014)
11. Popescu, A.C., Farid, H.: Exposing digital forgeries by detecting traces of resampling. IEEE Trans. Signal Process. 53(2), 758–767 (2005)
12. Cao, H., Kot, A.C.: Rich models for steganalysis of digital images. IEEE Trans. Inf. Forensics Secur. 7(3), 992–1002 (2012)
13. Verdoliva, L., Cozzolino, D., Poggi, G.: A feature-based approach for image tampering detection and localization. In: 2014 IEEE International Workshop on Information Forensics and Security (WIFS), Atlanta (2014)
14. Qi, X., et al.: A universal image forensic strategy based on steganalytic model. In: Proceedings of the 2nd ACM Workshop on Information Hiding and Multimedia Security, Szalburg (2014)
15. Fan, W., Wang, K., Cayre, F.: General-purpose image forensics using patch likelihood under image statistical models. In: 2015 IEEE International Workshop on Information Forensics and Security (WIFS), Rome (2015)
16. Cao, G., et al.: Contrast enhancement-based forensics in digital images. IEEE Trans. Inf. Forensics Secur. 9(3), 515–525 (2014)
17. Popescu, A.C., Farid, H.: Exposing digital forgeries by detecting traces of resampling. IEEE Trans. Signal Process. 53(2), 758–767 (2005)
18. Cao, H., Kot, A.C.: Rich models for steganalysis of digital images. IEEE Trans. Inf. Forensics Secur. 7(3), 992–1002 (2012)
19. Qi, X., et al.: A universal image forensic strategy based on steganalytic model. In: Proceedings of the 2nd ACM Workshop on Information Hiding and Multimedia Security, Szalburg (2014)
20. Fan, W., Wang, K., Cayre, F.: General-purpose image forensics using patch likelihood under image statistical models. In: 2015 IEEE International Workshop on Information Forensics and Security (WIFS), Rome (2015)
21. Cozzolino, D., Poggi, G., Verdoliva, L.: Recasting residual-based local descriptors as convolutional neural networks: an application to image forgery detection. http://arxiv.org/abs/1703.04615 (2017). Accessed November 15, 2019
22. Chen, B., Li, H., Luo, W: Image processing operations identification via convolutional neural network. http://arxiv.org/abs/1709.02908 (2017). Accessed November 15, 2019
20. Tang, H., et al.: Detection of various image operations based on CNN. In: 2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), Kuala Lumpur (2017)
21. Li, H., et al.: Identification of various image operations using residual-based features. IEEE Trans. Circuits Syst. Video Techn. 28(1), 31–45 (2018)
22. Bayar, B., Stamm, M.C.: Constrained convolutional neural networks: a new approach towards general purpose image manipulation detection. IEEE Trans. Inf. Forensics Secur. 13, 2691–2706 (2018)
23. Pevný, T., Bas, P., Fridrich, J.: Steganalysis by subtractive pixel adjacency matrix. In: Proceedings of the 11th ACM Workshop on Multimedia and Security, Princeton (2009).
24. Fridrich, J., Kodovsky, J.: Rich models for steganalysis of digital images. IEEE Trans. Inf. Forensics Secur. 7(3), 868–882 (2012)
25. Shi, Y.Q., Suthiwan, P., Chen, L.: Textural features for steganalysis. In: Pattern Recognition and Machine Intelligence, 8th International Conference, Tezpur (2019)
26. Mazumdar, A., et al.: Detection of image manipulations using siamese convolutional neural networks. In: Proceedings of the 6th ACM Workshop on Information Hiding and Multimedia Security, Innsbruck (2018)
27. Chen, Y., et al.: Densely connected convolutional neural network for multi-purpose image forensics under anti-forensic attacks. In: Proceedings of the 14th International Conference on Information Hiding, Berkeley (2013)
28. Boroumand, M., Fridrich, J.: Deep learning for detecting processing history of images. Electron. Imaging 2018, 213-1–213-9 (2018)
29. Chen, Y., et al.: A multi-purpose image forensic method using densely connected convolutional neural networks. J. Real-Time Image Process 16(3), 725–740 (2019)
30. Chen, Y., et al.: Depthwise separable convolutional neural network for image forensics, IEEE Visual Communications and Image Processing (VCIP), Sydney (2019)
31. Peng, A., et al.: A universal image forensics of smoothing filtering. Int. J. Digital Crime Forensics 11, 18–28 (2019)
32. Wei, Y., et al.: Auto-generating neural networks with reinforcement learning for multi-purpose image forensics. In: 2020 IEEE International Conference on Multimedia and Expo (ICME), London (2020)
33. Darimet, L., Wang, K., Cayre, F.: Graft: Unsupervised adaptation to resizing for detection of image manipulation. IEEE Access 8, 55619–55632 (2020)
34. Huang, J., Mumford, D.: Statistics of natural images and models. In: Proceedings of 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR000149), Fort Collins (1999)
35. Li, H., et al.: Image forgery localization via integrating tampering possibility maps. IEEE Trans. Inf. Forensics Secur. 12, 1240–1252 (2017)
36. Wang, J., et al.: Detection of image region duplication forgery using model with circle block. In: 2009 International Conference on Multimedia Information Networking and Security, Wuhan (2009)
37. Pan, X., Lyu, S.: Region duplication detection using image feature matching. IEEE Trans. Inf. Forensics Secur. 5(4), 857–867 (2010)
38. Irene-Amerini, R.C., Rudy-Becarelli, Casini, M.: A feature-based forensic procedure for splicing forgery detection. Math. Probl. Eng. 2015, 6 (2015)
39. Liu, Y., et al.: Image forgery localization based on multi-scale convolutional neural networks. https://arxiv.org/abs/1706.07842 (2017). Accessed November 15, 2019
40. Cozzolino, D., Gragnaniello, D., Verdoliva, L.: Image forgery localization through the fusion of camera-based, feature-based and pixel-based techniques. In: 2014 IEEE International Conference on Image Processing (ICIP), Paris (2014)
41. Yang, J., et al.: Detecting median filtering via two-dimensional ar models of multiple filtered residuals. Multimedia Tools and Applications 77(7), 7931–7953 (2018)
42. Yuan, H.D.: Blind forensics of median filtering in digital images. IEEE Trans. Inf. Forensics Secur. 6(4), 1335–1345 (2011)
43. Bas, P., Furon, T.: Bows-2. http://bows2.ec-lille.fr/ (2007). Accessed May 2020
44. IEEE IFS-TC image forensics challenge. http://ifc.recod.ic.unicamp.br/fc.webiste/index.py?sec=5. Accessed March 1, 2019
45. Kodovsky, J., Fridrich, J.: Steganalysis of JPEG images using rich models. Proc. SPIE 8303, 81–93 (2012).
46. Bas, P., Filler, T., Pevný, T.: Break our steganographic system: The ins and outs of organizing boss. In: Information Hiding 13th International Conference, Prague (2011)
47. Image dataset for localization of double jpeg compression. https://drive.google.com/drive/folders/1CdJ5QUdgZMCA8DEj0WCBgr2C__Y670n6k (2020). Accessed March 2, 2021
48. Kodovsky, J., Fridrich, J., Holub, V.: Ensemble classifiers for steganalysis of digital media. IEEE Trans. Inf. Forensics Secur. 7(2), 432–444 (2012)