Received: 20 May 2024 | Revised: 27 February 2025 | Accepted: 28 March 2025 | Published online: 14 May 2025

RESEARCH ARTICLE

Artificial Intelligence and Applications
2025, Vol. 00(00) 1–12
DOI: 10.47852/bonviewAIA52023454

Robust Source Camera Identification of Image Using Mean Subtracted Contrast Normalization for Digital Forensic Investigation

Pabitra Roy[1], Shyamali Mitra[2] and Nibaran Das[3,*]

[1]*Department of Computer Science and Engineering, Ramkrishna Mahato Government Engineering College, India*

[2]*Department of Instrumentation & Electronics Engineering, Jadavpur University, India*

[3]*Department of Computer Science and Engineering, Jadavpur University, India*

Abstract: This paper aims to investigate source camera identification (SCI), one of the challenges in image forensics. Besides SCI, research on the robustness of the SCI algorithm for practical applications is necessary because images are altered due to JPEG compression, Gaussian white noise, and rescaling on social networking platforms, where the fingerprints of the images may be contaminated. In this study, we explore robust SCI by extracting sensor pattern noise (SPN) for each camera model using mean subtracted contrast normalization (MSCN). Firstly, MSCN is extracted for every camera model. In this study, it is termed basic sensor pattern noise (BSPN). We further enhance the basic sensor pattern in the Fourier domain to obtain the final fingerprint, termed SPN. To attribute an unknown image to its source camera, the SPN of the image is extracted. The SPN of the unknown image is then correlated with the reference SPN of all camera models, respectively. If the correlation is greater than the particular threshold, then it leads to the camera model identification. Experiment results confirm that the proposed method effectively attributes an unknown image with its source camera and can resist JPEG compression, Gaussian white noise, and rescaling attacks more efficiently than the state-of-the-art SCI methods. Furthermore, the time required to extract SPN from the query image is also low.

Keywords: blind image forensic, mean subtracted contrast normalization, correlation, source camera identification, sensor pattern noise

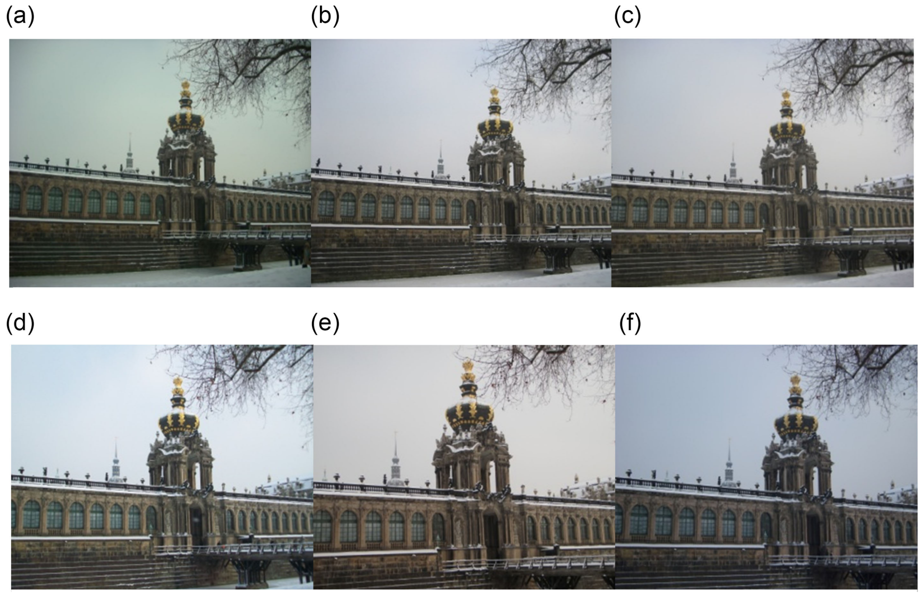
1. Introduction

The increasing digital crime in various forms has given rise to a wide research area in computer forensics that has several applications in digital forensic investigations. Currently, a variety of image-capturing devices enable the capture, storage, and daily posting of millions of images for public viewing. The purpose of camera model identification is to provide answers to the following questions: 1. “Which model of camera does this image (most likely) come from?” 2. “Is this photograph taken with a camera of this make and model?” [1]. The same image can be captured by different camera models as shown in Figure 1, and the problem is to identify the camera model for an image blindly. The effectiveness of image editing software is also expanding in parallel, allowing anonymous to forge digital image content. Anonymous may create phony accounts on social media or image content may be changed by using photo editing software. As a result, the actual owner of the image may be harassed by Law enforcement agencies. The images can only be used as “legal evidence” or “trusted evidence” when they are correctly authenticated. Digital camera processing is quite sophisticated, and it differs depending on the camera type and manufacturer. Because of their actual authenticity, digital photos are fragile. In situations involving child pornography and movie piracy, insurance claims, and scientific fraud, establishing the provenance of the imagery offered as only crucial evidence in court requires reliable and authentic identification of the capture device. Source camera identification (SCI) is a blind technique [2] as there is no previous embedded information like a digital watermark or digital signature within the image. In the digital image forensic domain, an image has to be fabricated to its source camera in a completely blind way. Blind image forensics is a strategy that addresses the following two problems primarily: unauthorized modification of image and image source identification [3]. Images are inherently related to the capturing device. It inherits sensor features from capturing devices, and cameras of different brands and models have different features. Even the same model can have a slightly different feature because human operating errors and other errors affect the correctness of the sensor.

In this study, we have employed a simplified model to construct a less complex camera model identification approach adaptable to a wide range of cameras using MSCN coefficients. The proposed method is tested on the “Dresden Image Database”, and the results are promising in terms of accuracy, speed, and robustness. The paper is organized as follows: Section 2 gives related work. Section 3 provides SCI using MSCN with the proposed methodology in detail. Section 4 represents the experimental

*Corresponding author: Nibaran Das, Department of Computer Science and Engineering, Jadavpur University, India. Email: nibaran.das@jadavpuruniversity.in

Figure 1
Same image captured with different camera models (a) Agfa_DC-830i, (b) Canon_Ixus55, (c) Canon_Ixus70, (d) Kodak_M1063, (e) Olympus_mju_1050SW, (f) Samsung_L74wide



setup. Section 5 gives experimental results with robustness. Finally, Section 6 concludes the paper.

2. Related Works

The sensor pattern noise-based (SPN) SCI techniques, such as the pixel non-uniformity noise-based method, maximum likelihood estimator of Photo Response Non-Uniformity (PRNU), The BM3D filter, Content Adaptive Guided image filter, and Dual-Tree complex wavelet transform based, etc., have been described in [4–11]. Under ideal conditions, the true positive rate (TPR) for a False-positive rate (FPR) at 10^{-4} reaches a peak of 23.23 (DTCWT). While most methods exhibit moderate speed, CAGIF outperforms them all in terms of speed but compromises on TPR. However, robustness remains a challenge across all methods. A multiscale feature fusion network is proposed in [12] to enhance SCI by extracting and fusing SPN features from image patches of different scales. A semi-supervised ensemble learning method (multi-DS) is proposed in [13] for SCI. A novel method for video SCI using noise patterns and majority voting is presented in [14]. CNN-based SCI has been discussed in [15–18]. CNN-based methods demonstrated an impressive accuracy of 97%. However, they lagged behind traditional methods in speed due to their high computational demands. For evaluation accuracy; state-of-the-art machine learning algorithms like SVM, logistic regression, and random forest-based classification have been used in most of the papers. These methods delivered an impressive accuracy ranging from 90% to 97% while maintaining a moderate processing speed. To identify the source camera, modified CNN such as AlexNet and local binary pattern are used [19] to improve the accuracy. This paper offers a balanced trade-off, delivering moderate accuracy compared to standard CNN-based methods while significantly reducing training time. To enhance the reliability of camera identification, undesirable artifacts should be removed from SPN [20]. Though EXIF tags could be used for SCI, EXIF metadata changes over time [21]. In [22], wavelet-based feature extraction and image classification using SVM are presented. In [23], depth cameras and their noise patterns are discussed. The Forchheim Image Database has been proposed in [24] and allows to cleanly separate image content from forensic artifacts. A deep learning-based approach that learns the unique traces from the images transformed to the discrete cosine and wavelet domains is shown in [25]. The effects of off-nominal exposure are studied in [26]. This paper uses DWT denoising methods with signed peak-to-correlation energy (PCE) for correlation. In [27], the authors extract SPN with 3 channels using DWT methods and then enhance SPN using model 5, as described in [28]. This method offers fast processing and delivers satisfactory results under ideal conditions; however, it lacks resilience against compression, rescaling, and Gaussian noise. [29] provides a comprehensive review of SCI methodologies, offering an in-depth exploration of the field. An advanced ConvNet-based SCI approach is introduced in [30], showcasing the power of deep learning. Pixel analysis-driven SCI techniques are discussed in [31], while [32, 33] present innovative methods leveraging color correlation features and multiscale feature fusion for enhanced SCI performance. SCI and detection in digital videos through blind forensics have been introduced in [34]. Fixed pattern noise removal, leveraging a pre-calibrated noise pattern, is demonstrated effectively in [35]. PRNU-based video source attribution, a method that utilizes PRNU to trace the origin of video content, has been meticulously developed and demonstrated in [36]. Non-distortion-specific features, derived using normalized Discrete Cosine Transform coefficients and modeled through a Generalized Gaussian Distribution, have been proposed in [37] for the SCI.

Though state-of-the-art SCI methods provide good accuracy, most of the approaches suffer from low robustness and high complexity. A robust and simple system generally requires less effort to extract the fingerprint of the query image and to speed up the identification process. Besides accuracy, robustness methods are the main concern in today's socioeconomic scenario. So we emphasize our work on robustness in SCI. The performance of the proposed method is measured in terms of TPR for fixed FPR, AUC (area under the curve), and speed to extract the SPN of the query image. The contributions can be encapsulated as follows:

1) The detection of SPN of the camera models using the MSCN of red and green channels and its enhancement to correctly attribute the source camera of an image is discussed. Using the test image SPN of the red and green channel, correlation is calculated with the reference SPN of all camera models.
2) The proposed algorithm provides high robustness when images are attacked by JPEG compression, Gaussian white noise, and rescaling. The proposed framework outperforms some recent works in terms of speed and robustness. The method also provides comparable results with state-of-the-art SCI methods.

3. SCI Using MSCN

The image sensor is the heart of every digital camera. In a traditional digital camera, there are numerous processing phases. Various errors in the image acquisition process are unavoidable. The sensor is divided into pixels, which gather photons and convert them to voltages. Voltage amplification and quantization create the output digital value. Before reaching the sensor, light from the captured scene goes via the camera lenses, a filter, and a color filter array (CFA). At each pixel, the CFA permits only one color to be measured. The digitized sensor data are subsequently interpolated (demosaicked) using color interpolation methods to obtain the three fundamental colors for each pixel. The digital image is then recorded in the camera's internal memory.

3.1. Conventional methods to extract SPN

Due to silicon wafer inhomogeneity, the Sensor Pattern Noise (SPN) created by Photo-Response Non-Uniformity (PRNU) varies from one camera to another, even within the same model and make. As a result, the source camera can be effectively identified using sensor-based pattern noise, which is calculated using Equation (1).

$$W = Y - F(Y) \tag{1}$$

where F is the denoising filter, Y is the image, and W is the noise residue. The denoising filter is critical for extracting pure SPN. The SPN of a specific camera model k is calculated using Equation (2) by averaging a number of noise residues from the same camera model [38] as:

$$\mathrm{SPN}(k) = \frac{\sum_{i=1}^{N} W_i}{N} \tag{2}$$

or using Equation (3), the maximum likelihood approach [39] is calculated as:

$$\mathrm{SPN}(k) = \frac{\sum_{i=1}^{N} Y_i W_i}{\sum_{i=1}^{N} Y_i^2} \tag{3}$$

where Y_i, i$\in \{1, 2, 3, \ldots, N\}$ denotes N photos captured by the same camera model, W_i is the noise residue of the i^{th} image, $W_i = Y_i - Y_i^{(0)}$, $Y_i^{(0)}$, is Y_i denoised using a denoising filter. The number of images used to calculate SPN for camera model k is N. All the operations in Equation (3) are elementwise. Image content is suppressed in this fashion, and SPN is calculated correspondingly. For identification, normalized correlation coefficients, a statistical hypothesis test, and PCE are typically utilized.

3.2. Mean subtracted contrast normalization (MSCN)

In image processing and computer vision tasks, the MSCN technique is frequently employed. The main characteristics for which MSCN has been used to extract SPN are listed below. By normalizing the contrast, it keeps the structure of an image intact. A more informative image is produced by normalizing the local contrast, which enhances both low- and high-contrast areas of an image. When lighting conditions vary between or within images, MSCN is especially helpful. MSCN effectively removes the global illumination component from each pixel by subtracting the mean intensity from each pixel, making the image more invariant to lighting variations, which is an important aspect of SCI. This is useful for SCI, as images captured in different lighting conditions can have a significant impact on performance. The noise component is attenuated by subtracting the mean value, making the image more resistant to noise interference. MSCN is one of the modern techniques which transform the image intensity into luminance at a given pixel. The MSCN coefficients are calculated using the following steps, as illustrated in Figure 2.

On the log–contrast values, local mean subtraction and variance normalization follow the Gaussian distribution observed by Daniel L Ruderman [40]. MSCN coefficients are calculated using the following Equations (4), (5) and (6).

$$\hat{I}(i,j) = \frac{I(i,j) - \mu(i,j)}{\sigma(i,j) + c} \tag{4}$$

where $I(i, j)$ is the image intensity at a given pixel(i,j) and $\hat{I}(i,j)$ is the luminance corresponding to the pixel, where $i \in \{1, 2, 3, \ldots, H\}, j \in \{1, 2, 3, \ldots, W\}$ are the spatial indices (H and W are the height and width of the image, respectively), $\mu(i, j)$ is the local mean-field and $\sigma(i, j)$ is the local variance field.

$$\mu(i,j) = \sum_{m=-M}^{M} \sum_{n=-M}^{M} W_{m,n} I(i+m, j+n) \tag{5}$$

$$\sigma(i,j) = \sqrt{\sum_{m=-M}^{M} \sum_{n=-M}^{M} W_{m,n} [I(i+m, j+n) - \mu(i,j)]^2} \tag{6}$$

where $W = \{w_{m,n} | m = -M, \ldots, M, n = -N, \ldots, N\}$ is a 2D circularly symmetric Gaussian weighting function sampled out to 3 standard deviations ($M = N = 3$) and rescaled to unit volume. In Equation (4), μ is the denoised version of an image I with a Gaussian filter. The Gaussian kernel filter is designed with a window size of 7×7. The value of c = 1 is a constant in Equation (4) that is used to prevent instabilities when the denominator approaches zero. As the SPN is unique to the camera model, the MSCN is used here to extract the SPN in conjunction with Equation (2), described in Section 3.4 in more detail. In Figure 3, the observation of the MSCN coefficient is shown.

Figure 2
Steps to calculate MSCN coefficients

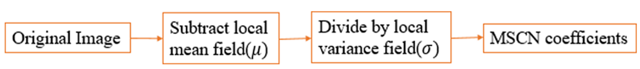
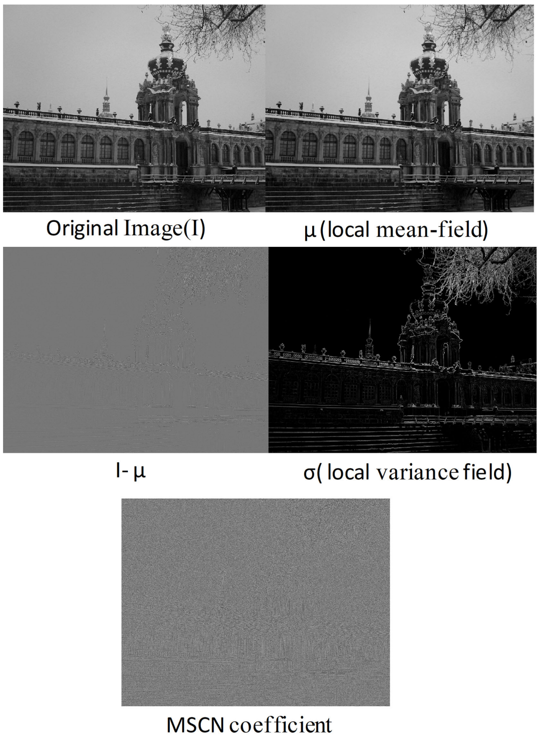


Figure 3
MSCN coefficient calculation



3.3. The proposed methodology

The overview of the proposed method is shown in Figure 4. The MSCN coefficients of the green and red channels of a number of images are used to determine the basic sensor pattern noise (BSPN). The reason for using green and red channels is that the Bayer pattern has twice as many green pixels in the matrix as red or blue pixels. When the RGB triplets are rebuilt from neighboring sensor pixels, the blue and red channels immediately fall behind the green channel, resulting in significantly higher spectral noise for those two channels [41]. Second, current CMOS sensors are around 50% more sensitive to the green and red regions of the spectrum than to the blue. Blue has a significantly worse signal-to-noise ratio than the other two colors due to its lack of sensitivity combined with a fixed sensor level and sampling noise for all pixels throughout the sensors. In experiments to determine the best channel, it is found that the red and green channels work better together than the red, green, blue, red-green, blue-green, and green-blue channels. By including the blue channel along with the other two channels, the performance remains the same. To decrease the running time to extract SPN from the query image, the blue channel is disregarded to enable the proposed method to be used in real time. Figure 5 illustrates the channel comparison in terms of TPR for fixed FPR 10^{-4}.

3.4. Extraction of SPN

In order to extract SPN from an image, BSPN is extracted using algorithm 1, Equation (2), and the method described in [38] with red and green channel MSCN coefficients. During the estimation of SPN, the extracted signal contains additional non-unique components, such as artifacts due to color interpolation, on-sensor signal transfer, and JPEG compression (blockiness). These artifacts are not unique to a specific camera sensor but are shared among cameras of the same brand or those using similar sensor designs. This may introduce correlations between SPNs of different cameras, leading to an increase in false identification rates and a decrease in the reliability of the camera identification process. To mitigate this issue, these non-unique artifacts are removed using Fourier domain enhancement of SPN. Hence, discrete Fourier domain is adapted as used by [5] in algorithm 2. First DFT of the BSPN is computed and

Figure 4
Overview of the proposed method using MSCN coefficients

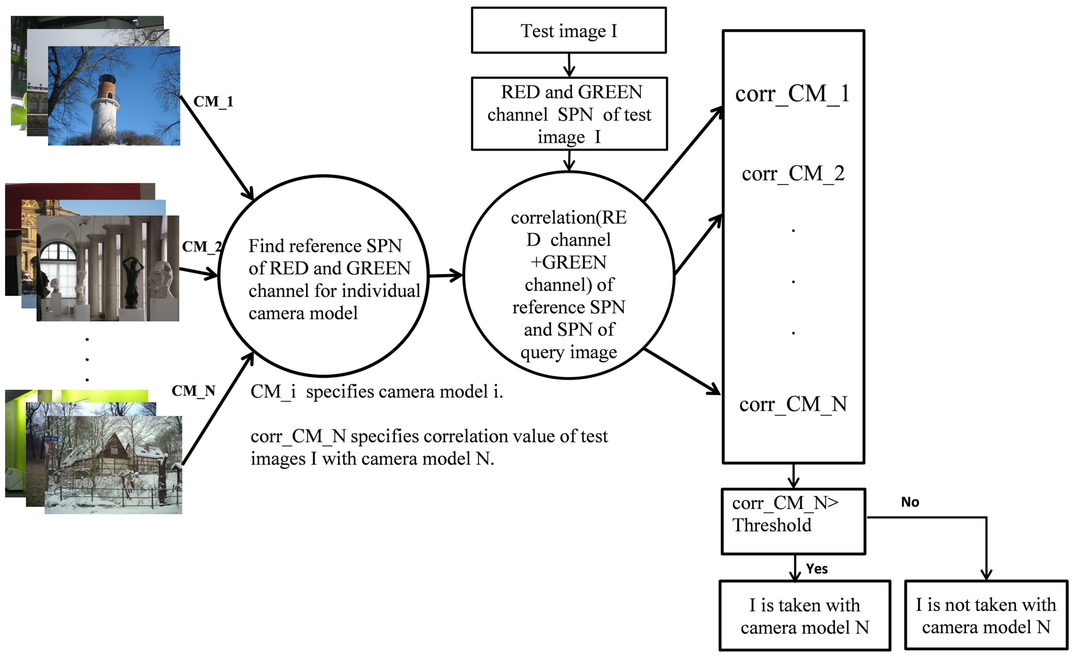
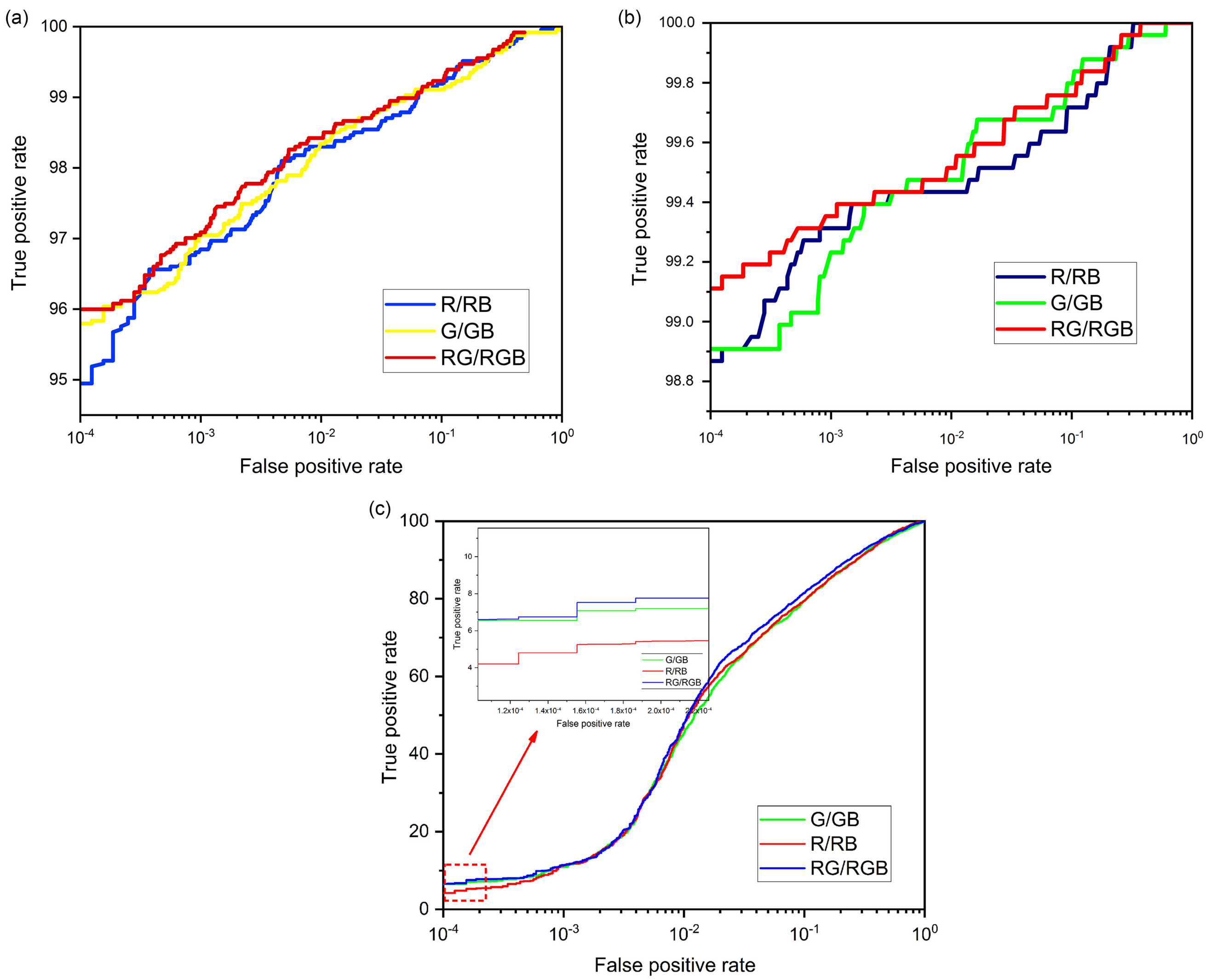


Figure 5
Performance comparison of RGB channels in terms of ROC curves
(a) Medium size (1024 × 768), (b) big size (2592 × 1944), (c) small size (128 × 128)



takes the real part only. Estimate the noise variance of the BSPN in the frequency domain using MAP (maximum a posteriori probability) estimation for two sizes of square N x N neighborhoods [N = 3, 5]. The final estimate is a minimum of two. The noise coefficient is then calculated, and a Wiener filter-like attenuation is applied to obtain denoised BSPN in the frequency domain. By taking the inverse DFT, pure SPN is extracted.

When a fingerprint is calculated, a non-unique artifact of an image that came from several cameras of the same model revealed only a modest similarity. This non-unique artifact is not exclusive to the sensor, and as a result, they raise the FPR. These non-unique artifacts and the image content need to be suppressed in order to recover pure SPN. Fortunately, the majority of these artifacts are caused by demosaicking methods, which are periodic in nature and rely on the CFA [42]. Along with these periodic artifacts, non-periodic artifacts (e.g., compression artifacts or artifacts inherent to sensor on-board circuitry) are also present in BSPN. To suppress periodic and non-periodic artifacts from BSPN, Algorithm 2 is introduced, as done in [1]. Step 1 in algorithm 2 is used to convert BSPN into the frequency domain, as the non-unique artifacts and image details from the scene are largely contaminated in the frequency domain. In step 2, the noise variance of the BSPN is calculated as the noise that we want to not exceed locally. Step 3 entails estimating the BSPN coefficient variance using MAP and noise variance with two sizes of square NxN neighborhoods [N = 3, 5] to avoid exceeding the noise variance locally. Equation (13) in algorithm 2 is used to extract the noise coefficient after suppressing non-unique artifacts present in BSPN.

3.5. Detection of camera model

The reference SPN of different camera models is correlated with the extracted SPN of the query image using both the red and green channels using Equations (7) and (8). The total correlation is calculated using Equation (9). If the total correlation value is greater than a particular threshold value, then it is decided that the query image is taken with that camera model; otherwise, the query image is not taken with that camera model. The steps are given below:

Steps:

1) Calculate the correlation of red and green channel SPN of query image with red and green channel reference SPN of camera model say A, using Equations (7) and (8).

$$corr\_red(SPN_{Red}, P_{c\_Red}) = \frac{\sum(SPN_{Red}.P_{c\_Red})}{\sqrt{\sum SPN_{Red}^2 \cdot \sum P_{c\_Red}^2}} \quad (7)$$

$$corr\_green(SPN_{Green}, P_{c\_Green}) = \frac{\sum(SPN_{Green}.P_{c\_Green})}{\sqrt{\sum SPN_{Green}^2 \cdot \sum P_{c\_Green}^2}} \quad (8)$$

where P_{c_Red} and P_{c_Green} are the red and green channel SPN for query image P respectively.

2) Calculate the total correlation using Equation (9) for all camera models, to decide whether image P has been captured by camera model A.

$$corr_T = corr\_red + corr\_green \quad (9)$$

where $corr_T$ is the total correlation value.

3) If the $corr_T$ value is greater than a particular threshold value, then P is taken with the camera model A, else P is not taken with camera model A. To maintain the recommended FPR, the detector threshold should be changed correspondingly.

4). Experimental Setup

In this experiment, images are taken from the "Dresden Image Database", a widely used benchmark dataset [43]. From this data set, a total of 14 camera models and devices were selected to carry out the experiments. For each camera model and device, 50 natural images are chosen at random from the dataset as training images to determine the reference SPN of each red and green channel of a specific camera model. The reason to choose RGB color space is that the Bayer filter is a very common RGB filter, and most camera models use it. Other CFAs, like RYYB CFA, have been introduced since 2019 and are used in Huawei's p30 series smartphones. X-trans CFA has been introduced in 2013, and it is only Fujifilm Camera-specific. We have not explored other CFA in this study and that will be our future work. Additionally, images are taken to test the proposed method of each model and device that are not in the training images.

Algorithm 1: To Find Basic Sensor Pattern Noise (BSPN)

Input: Color image
Output: Basic Sensor Pattern Noise (BSPN)

1) Find MSCN coefficients for each image using Equations (4), (5), and (6).
2) Extract the red and green channel MSCN coefficients.
3) Find the BSPN pattern of the red and green channels for a particular camera model using the following Equations (10) and (11):

$$BSPN_{Red}^c = \frac{\sum_{i=1}^N MSCN\_R_i^c}{N} \quad (10)$$

$$BSPN_{Green}^c = \frac{\sum_{i=1}^N MSCN\_G_i^c}{N} \quad (11)$$

where $MSCN\_R_i^c$ and $MSCN\_G_i^c$ represent the MSCN coefficients of the red and green channels for image i of camera model C, respectively. N is the number of images for a particular camera model C. $BSPN_{Red}^c$ and $BSPN_{Green}^c$ represent the BSPN of the red and green channels of camera model C, respectively.

4) Repeat steps 1 to 3 to find the BSPN of the red and green channels for all camera models.

Table 1 represents the experimental data set for the experiment, where the last digit of each camera model specifies the different devices of the same camera model. Three image sizes are tested in the study when the images are not subjected to any attack. For example, small (128 × 128), medium (1024 × 768), and large (2592 × 1944) images are used. To test the robustness, an image size of 128 × 128 is taken from the center of the image, which is the most commonly used patch size for SCI methods. Once one camera is designated as the reference camera, the images taken with this camera are treated as positive samples, whereas the images from all the other cameras are used as negative samples.

For example, the Canon Ixus55_0 camera model has 136 positive samples and (2474-136) = 2338 negative samples. A total of 2474 positive samples and 32162 negative samples are used in the experiment. The proposed method is compared with the state-of-the-art SCI methods such as DWT [39], CAGIF method [5], BM3D methods [4], DTCWT method [6], KLD method [7], and JSD method [11]. The baseline DWT method is used in the comparison. The CAGIF method with $\sigma = 5$ and smoothing parameter $\in = 0.02$ is used as indicated in [5]. As indicated in [6], the DTCWT method with a decomposition level l = 4 and standard deviation of SPN $\sigma = 1.8$ as recommended in [38, 39] is used. In the BM3D technique, σ is set 5. In [27], DWT with enhanced SPN, as described in [28], is used, where enhanced model 5 with α is set to 7.

Table 1
List of experiment camera models and image data sets

Sl. No.	Camera model	Number of images to find reference SPN	Number of images for testing
1	Agfa_DC-830i_0	50	287
2	Canon_Ixus55_0	50	136
3	Casio_EX-Z150_0	50	120
4	FujiFilm_FinePixJ50_0	50	160
5	Kodak_M1063_0	50	302
6	Nikon_CoolPixS710_0	50	127
7	Olympus_mju_1050SW_0	50	137
8	Olympus_mju_1050SW_1	50	134
9	Panasonic_DMC-FZ50_0	50	157
10	Panasonic_DMC-FZ50_1	50	312
11	Pentax_OptioA40_0	50	117
12	Ricoh_GX100_0	50	139
13	Samsung_L74wide_0	50	178
14	Samsung_L74wide_1	50	168

Algorithm 2: To Enhance the Basic Sensor Pattern Noise (BSPN)

Input: BSPN (Basic Sensor Pattern Noise) and standard deviation of BSPN (sigma)
Output: Enhanced SPN

1) Compute the discrete Fourier transform (DFT) of BSPN and take the real part, i.e.,

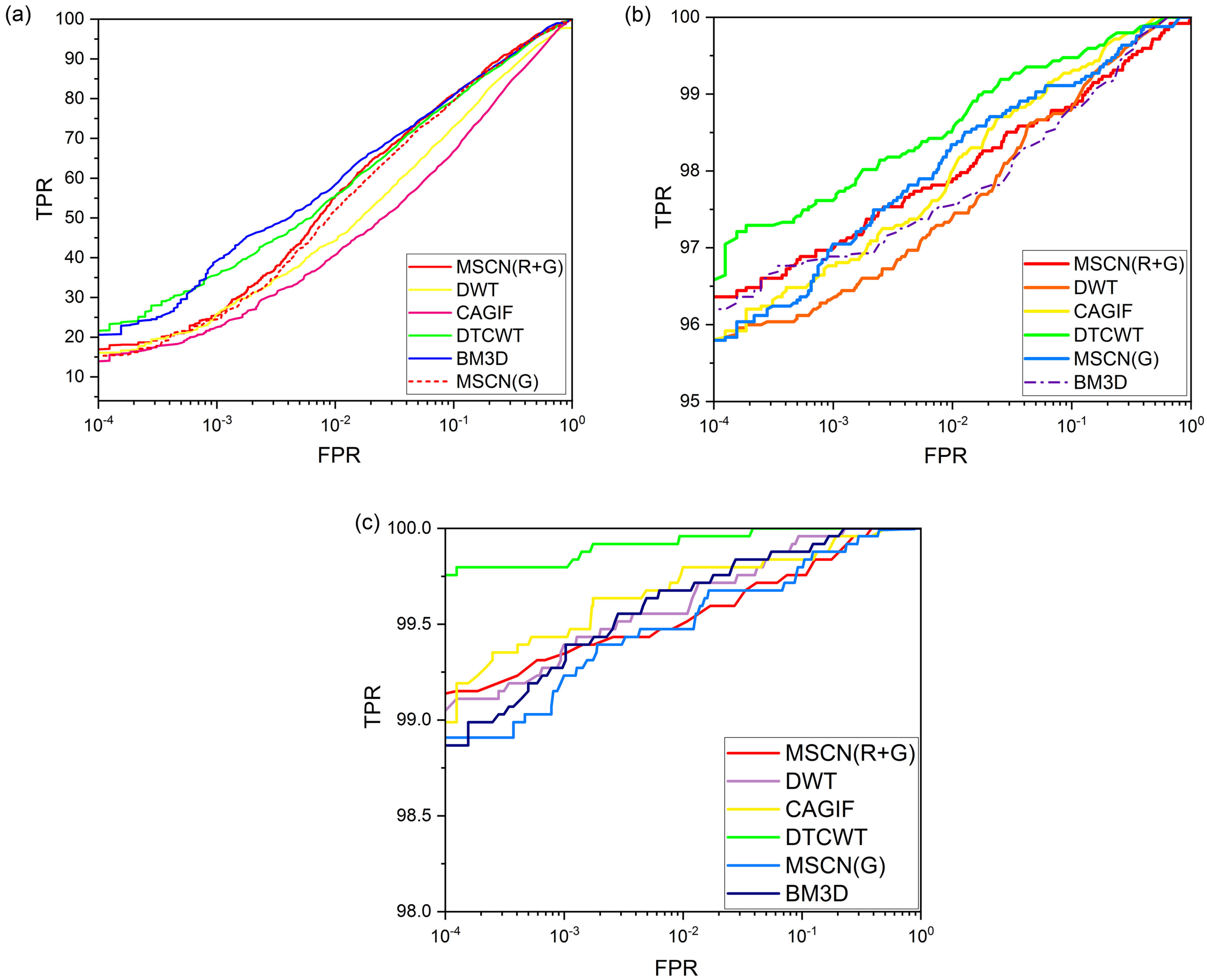
$$F = DFT(BSPN),\ D = real(F)$$

2) Estimate the variance: $variance = sigma^2$
3) Estimate coefficient variance (coeff) of D using maximum a posteriori probability (MAP) estimation for two sizes of square NxN neighborhoods [N = 3, 5].
4) Extract noise coefficient (N_c) using the Equation (12):

$$N_c = \frac{D.variance}{(coeff + variance)} \quad (12)$$

5) Compute F' (denoised version of BSPN in the frequency domain).

Figure 6
Performance comparison of different SCI methods in terms of ROC curves. From top to bottom: (a) small size (128 × 128), (b) medium size (1024 × 768), (c) big size (2592 × 1944), TPR = True-positive rate, FPR = False-positive rate



$$F' = \frac{F.N_c}{D} \tag{13}$$

6) Compute the inverse DFT of $F^{'}$ to get the final SPN.

$$\mathrm{SPN} = \mathrm{real}(\mathrm{IDFT}(F'))$$

5. Experimental Results

The true positive (TP) and false positive (FP) of each camera are determined for a specified detection threshold, and then, the overall TP and FP can be obtained. The TPR and FPR are defined as Equations (14) and (15):

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{14}$$

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}} \tag{15}$$

where FN and TN are the false negative and true negative, respectively. For normal images, DTCWT is shown to be the most accurate of the eight approaches for images of size 128 × 128, while our method is ranked third in terms of accuracy for fixed FPR 10^{-4}. The ROC curves in Figure 6 demonstrate a direct comparison of different techniques (Best Six). The FPR is on the x-axis, and the TPR is on the y-axis. For image size 128 × 128

Table 2
Average time comparison for extraction SPN from a query image (time in seconds)

Image size Method	1024 × 768 Running time(s)	2592 × 1944 Running time(s)
CAGIF [5]	0.52	2.92
DWT [39]	0.62	3.72
DTCWT [6]	0.78	5.36
MSCN(R+G)	0.6	3.01
BM3D [4]	4.81	29.09
MSCN(G)	0.43	1.99
KLD [7]	0.63	3.86
JSD [11]	0.64	3.99
ESPN [27]	0.75	3.96

Table 3
Average time comparison for extraction of SPN from a query image of size 500 × 500 (time in seconds)

Image size Method	500 × 500 Running time(s)
CAGIF [5]	0.12
DWT [39]	0.19
DTCWT [6]	0.23
MSCN (R+G)	0.13
BM3D [4]	2.16
KLD [7]	0.19
JSD [11]	0.20
ESPN [27]	0.22

at fixed FPR 10^{-4}, the TPR of eight methods are 16.2 (MSCN(R+G)), 14.2 (CAGIF), 16.0 (DWT), 20.1 (BM3D), 23.32 (DTCWT), 1.14(KLD), 13.06 (ESPN), and 1.45 (JSD), respectively. For image sizes of 1024 × 768 and 2592 × 1944, the proposed method ranked second for fixed FPR at 10^{-4}.

In a real-time application, the SPN extraction method from the query image is the main computational load of SCI. To provide an obvious speed comparison, the SPN for query images is extracted using several approaches from images of sizes 1024 × 768, 2592 × 1944, and 500 × 500 pixels. The average run times are compared in Tables 2 and 3, respectively. The simulations are performed with a MATLAB 2019a laptop with a 3.30 GHz Intel core i5 CPU and 4GB RAM. If the red and green channels are considered, the CAGIF approach is the fastest of the mentioned methods, while the MSCN(R+G) method comes in second for images of size 1024 × 768, 2592 × 1944 as shown in time comparison Table 2. In terms of the red and green channels, the MSCN(R+G)-based technique outperforms DWT, DTCWT, and BM3D methods, respectively. Table 3 shows the average run time for a 500 × 500 pixels image. KLD [7] employs 5 camera models for PRNU extraction, each of which requires a different amount of time to process.

To make a fair comparison across different methods, the average processing time for a 500 × 500 pixel-sized image is taken. The recommended approach came in second place here as well. The limitation of the proposed method is that for almost dark images, the correlation value is very poor. As a result, the false negative increases, as shown in Figure 7. The TPR for a given FPR of 10^{-4} varies when a random number of images are used to generate the reference SPN in increasing order, as illustrated in Figure 8. The proposed method indicates that the performance is optimum when the number of images taken to find the reference SPN is 80.

5.1. Robustness of proposed method

For speed and storage, most images on the internet are JPEG compressed, and it is well known that when an image is JPEG compressed, the SPN suffers. Similarly, when an image is subjected to Gaussian noise and rescaling, the SPN is weakened. The test images are JPEG compressed with a Quality factor of 75 and 90, rescaled with a factor of 0.9 and 1.1, and Gaussian attacked with a mean of 0 and variance of 0.01 and 0.001, respectively. Despite modest ROC performance degradation, all of the approaches are resilient to JPEG compression, rescaling, and Gaussian noise. The MSCN approach outperforms the seven methods. Figure 9 shows the direct comparison in terms of robustness including JPEG compression (QF = 75), rescaling (0.9), and Gaussian attack (0.01). Only the best six methods are shown in graphical representation. Also, it is observed that none of the methods provides the same performance with respect to the three attacks. For example, the CAGIF method is more robust in Gaussian noise than the rescaling attack. The proposed method provides the same efficiency in terms of robustness concerning JPEG compression, rescaling, and Gaussian noise. In all cases, the image is cropped from the center of the image to a size of 128 × 128. The numerical results are shown in Table 4 of image size 128 × 128. The AUC of different methods is listed in Table 5.

Figure 7
Few image samples which are incorrectly identified by our proposed method

Figure 8
Performance of MSCN(R+G) in terms of ROC curves by taking a random number of images in increasing order to generate reference SPN. Image size is taken 1024 × 768, TPR = True-positive rate, FPR = False-positive rate

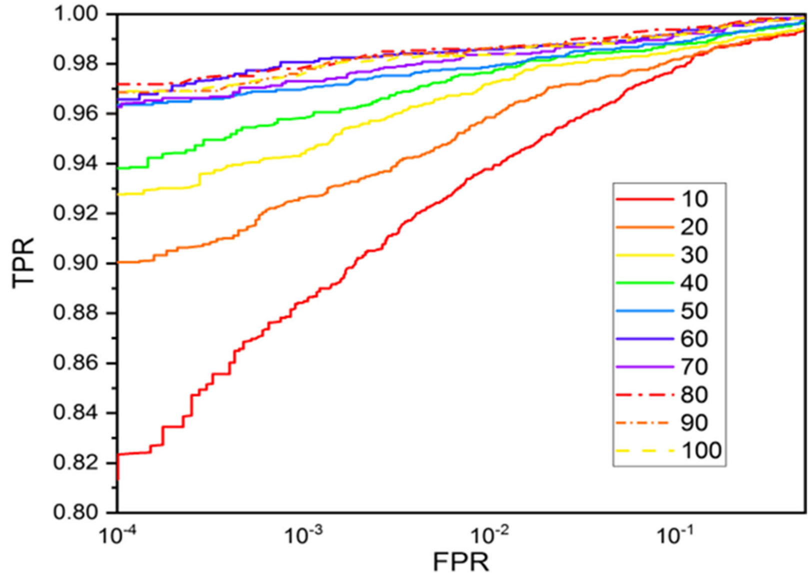


Figure 9
Comparison of robustness from top to bottom: (a) JPEG compression (QF = 75), (b) rescaling (0.9), and (c) Gaussian noise (mean = 0, variance = 0.01)

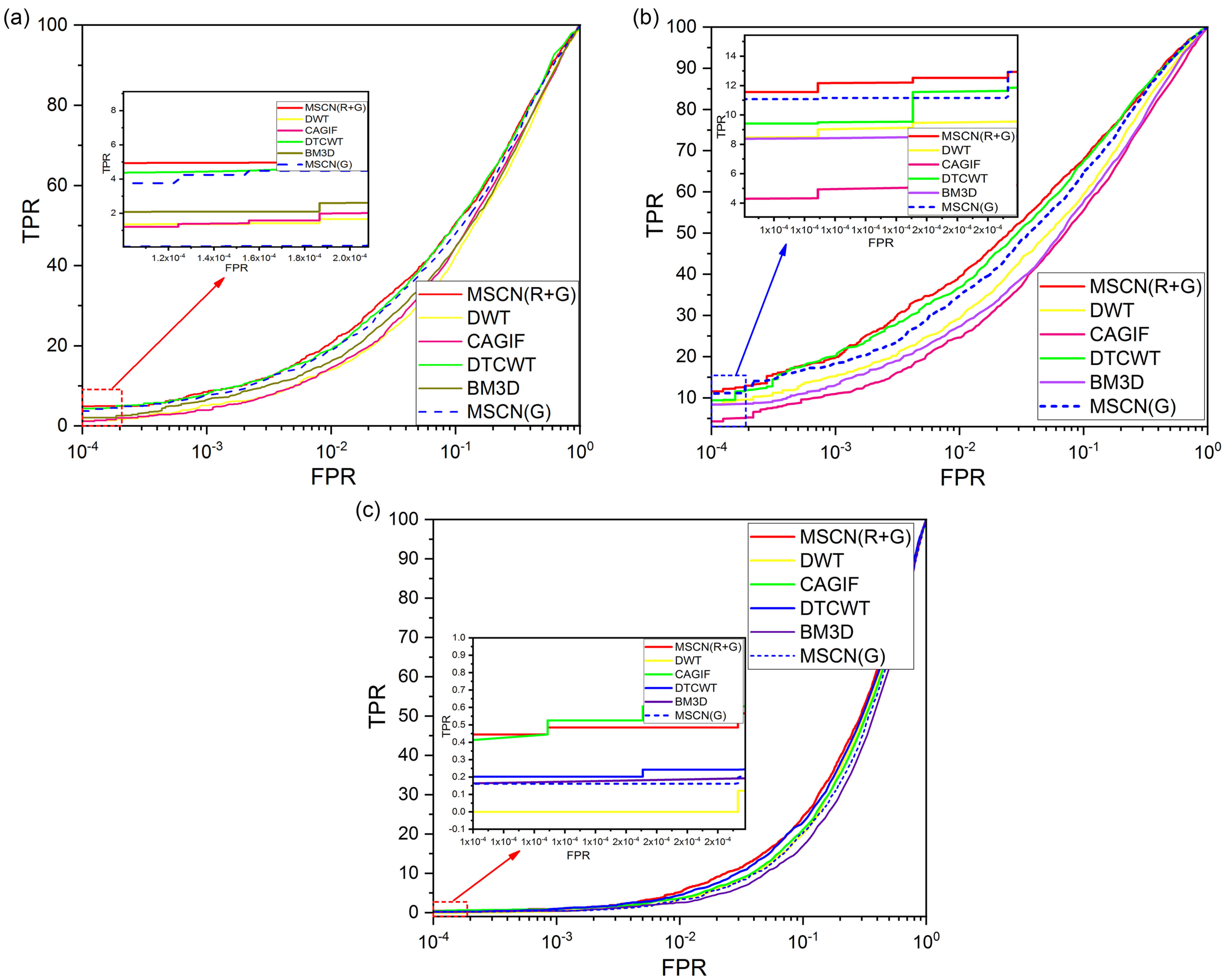


Table 4
TPR of different attacks for FPR at 10^{-4}. For each attack, the best result appears in bold font

Attack	MSCN(R+G)	DWT	CAGIF	BM3D	DTCWT	KLD	JSD	ESPN
JPEG compression (QF = 75)	**4.94**	1.37	1.21	2.1	4.4	0.54	0.61	1.33
JPEG compression (QF = 90)	**8.81**	2.12	1.89	4.75	5.09	1.12	1.2	1.88
Rescaling (0.9)	**12.66**	9.01	4.93	8.4	9.49	0.78	0.58	4.8
Rescaling (1.1)	**11.50**	8.75	5.10	8.12	8.36	0.65	0.67	4.4
Gaussian noise (mean = 0, variance = 0.01)	**0.48**	0.04	0.44	0.2	0.2	0.05	0.06	0.36
Gaussian noise (mean = 0, variance = 0.001)	**2.18**	0.09	1.23	1.2	1.1	0.06	0.08	1.2

Table 5
Comparison of AUC (area under curve) for the image of size 128 × 128 for fixed FPR at 10^{-4}. The best result appears in bold font

Methods	Normal	Rescaling (0.9) attack	Gaussian noise attack (Mean = 0, Variance = 0.01)	JPEG compression attack (QF = 75)
MSCN(R+G)	91.50	**87.34**	**64.96**	**79.61**
DWT [39]	88.71	83.8	62.59	75.46
CAGIF [5]	86.71	82.85	62.44	76.15
DTCWT [6]	91.70	86.90	63.95	79.6
BM3D [4]	**92.45**	82.46	63.07	76.87
MSCN(G)	89.78	85.42	60.89	77.76
KLD [7]	83.60	80.52	62.33	72.66
JSD [11]	83.65	80.56	62.89	72.86
ESPN [27]	85.55	85.25	64.2	76.27

6. Conclusions

The SCI problem is re-examined in this study using the MSCN coefficient with accuracy, speed, and robustness. The newly proposed method uses MSCN coefficients to extract SPN from a query image. The proposed method is compared with other cutting-edge SCI techniques. When images are not modified, the proposed method produces comparable results, but when images are compressed, rescaled, and exposed to Gaussian noise, the proposed method produces the best results in terms of TPR for fixed FPR at 10^{-4} than the state-of-the-art methods. In terms of AUC, the proposed method is best for rescaling and Gaussian attack, but it ranked second for JEPG compression. Despite being the most robust of the seven approaches, the proposed method came in second-best after CAGIF in terms of speed when using both the red and green channels together. When images are not manipulated, however, the DTCWT and BM3D methods produce the best results, but the time to extract SPN is significant, and though the CAGIF approach is the fastest, its performance falls short of expectations. The proposed method, on the other hand, outperformed CAGIF. This advantage will allow the proposed method to be applied in real-time applications. However, for the almost dark images, the proposed method gives a poor value of correlation, and as a result, it increases the false negative. Handling this constraint and developing an SPN of MSCN-based strategy should be investigated in the future. Further, the Spectral Cross-Domain Neural Network [44] could be trained on images from various camera brands and models, learning their unique spectral signatures for SCI. The limitation of the proposed work may be overcome by brightening the image during the preprocessing stage. Furthermore, enhancing the SPN could improve the system's performance and may serve as future work under different ISO sensitivity conditions.

Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

The Dresden image database that supports the findings of this study is openly available at https://doi.org/10.1145/1774088.1774427, reference number [43].

Author Contribution Statement

Pabitra Roy: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization. **Shyamali Mitra:** Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision. **Nibaran Das:** Validation, Formal analysis, Resources, Data curation, Writing – review & editing, Visualization, Supervision.

References

[1] Verdoliva, L. (2016). Handbook of digital forensics of multimedia data and devices [book reviews]. *IEEE Signal Processing Magazine*, *33*(1), 164–165. https://doi.org/10.1109/MSP.2015.2488018

[2] Sameer, V. U., Sugumaran, S., & Naskar, R. (2018). *K*-unknown models detection through clustering in blind source camera identification. *IET Image Processing*, *12*(7), 1204–1213. https://doi.org/10.1049/iet-ipr.2017.1142

[3] Yao, H., Qiao, T., Xu, M., & Zheng, N. (2018). Robust multi-classifier for camera model identification based on convolution neural network. *IEEE Access*, *6*, 24973–24982. https://doi.org/10.1109/ACCESS.2018.2832066

[4] Cortiana, A., Conotter, V., Boato, G., & de Natale, F. G. B. (2011). Performance comparison of denoising filters for source camera identification. *Media Watermarking, Security, and Forensics: Proceedings of SPIE*, *7880*, 60–65.

[5] Zeng, H., & Kang, X. (2016). Fast source camera identification using content adaptive guided image filter. *Journal of Forensic Sciences*, *61*(2), 520–526. https://doi.org/10.1111/1556-4029.13017

[6] Zeng, H., Wan, Y., Deng, K., & Peng, A. (2020). Source camera identification with dual-tree complex wavelet transform. *IEEE Access*, *8*, 18874–18883. https://doi.org/10.1109/ACCESS.2020.2968855

[7] Quintanar-Reséndiz, A. L., Rodríguez-Santos, F., Pichardo-Méndez, J. L., Delgado-Gutiérrez, G., Ramírez, O. J., & Vázquez-Medina, R. (2021). Capture device identification from digital images using Kullback-Leibler divergence. *Multimedia Tools and Applications*, *80*(13), 19513–19538. https://doi.org/10.1007/s11042-021-10653-1

[8] Taspinar, S., Mohanty, M., & Memon, N. (2020). Camera identification of multi-format devices. *Pattern Recognition Letters*, *140*, 288–294. https://doi.org/10.1016/j.patrec.2020.10.010

[9] Iuliani, M., Fontani, M., & Piva, A. (2021). A leak in PRNU based source identification—Questioning fingerprint uniqueness. *IEEE Access*, *9*, 52455–52463. https://doi.org/10.1109/ACCESS.2021.3070478

[10] Bernacki, J., & Scherer, R. (2022). Digital forensics: A fast algorithm for a digital sensor identification. *Journal of Information and Telecommunication*, *6*(4), 399–419. https://doi.org/10.1080/24751839.2022.2058252

[11] Rodríguez-Santos, F., Quintanar-Reséndiz, A. L., Delgado-Gutiérrez, G., Palacios-Luengas, L., Jiménez-Ramírez, O., & Vázquez-Medina, R. (2022). Identifying the digital camera from natural images using residual noise and the Jensen–Shannon divergence. *Journal of Electrical and Computer Engineering*, *2022*(1), 1574024. https://doi.org/10.1155/2022/1574024

[12] Hui, C., Jiang, F., Liu, S., & Zhao, D. (2022). Source camera identification with multi-scale feature fusion network. In *2022 IEEE International Conference on Multimedia and Expo*, 1–6. https://doi.org/10.1109/ICME52920.2022.9859965

[13] Wang, B., Hou, J., Ma, Y., Wang, F., & Wei, F. (2022). Multi-DS strategy for source camera identification in few-shot sample data sets. *Security and Communication Networks*, *2022*(1), 8716884. https://doi.org/10.1155/2022/8716884

[14] Timmerman, D., Bennabhaktula, S., Alegre, E., & Azzopardi, G. (2020). Video camera identification from sensor pattern noise with a constrained ConvNet. *arXiv Preprint: 2012.06277*.

[15] Anjankar, S., Telang, S., Bharadwaj, K., & Khandelwal, R. (2022). Identification of source camera by amalgamation of PRNU and noise print using dimensionality expansive residual network. *International Journal of Next-Generation Computing*, *13*(5), 1172. http://dx.doi.org/10.47164/ijngc.v13i5.919

[16] Liu, Y., Zou, Z., Yang, Y., Law, N. F. B., & Bharath, A. A. (2021). Efficient source camera identification with diversity-enhanced patch selection and deep residual prediction. *Sensors*, *21*(14), 4701. https://doi.org/10.3390/s21144701

[17] Manisha, Li, C. T., Lin, X., & Kotegar, K. A. (2022). Beyond PRNU: Learning robust device-specific fingerprint for source camera identification. *Sensors*, *22*(20), 7871. https://doi.org/10.3390/s22207871

[18] Manisha, Li, C. T., & Kotegar, K. A. (2023). Source camera identification with a robust device fingerprint: Evolution from image-based to video-based approaches. *Sensors*, *23*(17), 7385. https://doi.org/10.3390/s23177385

[19] Wang, B., Yin, J., Tan, S., Li, Y., & Li, M. (2018). Source camera model identification based on convolutional neural networks with local binary patterns coding. *Signal Processing: Image Communication*, *68*, 162–168. https://doi.org/10.1016/j.image.2018.08.001

[20] Gupta, B., & Tiwari, M. (2018). Improving performance of source camera identification by suppressing peaks and eliminating low-frequency defects of reference SPN. *IEEE Signal Processing Letters*, *25*(9), 1340–1343. https://doi.org/10.1109/LSP.2018.2857223

[21] Mullan, P., Riess, C., & Freiling, F. (2019). Forensic source identification using JPEG image headers: The case of smartphones. *Digital Investigation*, *28*, S68–S76. https://doi.org/10.1016/j.diin.2019.01.016

[22] Sandoval Orozco, A. L., García Villalba, L. J., Arenas González, D. M., Corripio, J. R., Hernandez-Castro, J., & Gibson, S. J. (2015). Smartphone image acquisition forensics using sensor fingerprint. *IET Computer Vision*, *9*(5), 723–731. https://doi.org/10.1049/iet-cvi.2014.0243

[23] Haider, A., & Hel-Or, H. (2022). What can we learn from depth camera sensor noise?. *Sensors*, *22*(14), 5448. https://doi.org/10.3390/s22145448

[24] Hadwiger, B., & Riess, C. (2021). The Forchheim image database for camera identification in the wild. In *International Conference on Pattern Recognition*, 500–515. https://doi.org/10.1007/978-3-030-68780-9_40

[25] Manisha, Karunakar, A. K., & Li, C. T. (2021). Identification of source social network of digital images using deep neural network. *Pattern Recognition Letters*, *150*, 17–25. https://doi.org/10.1016/j.patrec.2021.06.019

[26] Martin, A., Maxion, R., & Newman, J. (2024). Forensic camera identification: Effects of off-nominal exposures. *arXiv Preprint: 2407.00543*.

[27] Li, Q., Wang, R., & Xu, D. (2023). A video splicing forgery detection and localization algorithm based on sensor pattern noise. *Electronics*, *12*(6), 1362. https://doi.org/10.3390/electronics12061362

[28] Li, C. T. (2010). Source camera identification using enhanced sensor pattern noise. *IEEE Transactions on Information Forensics and Security*, *5*(2), 280–287. https://doi.org/10.1109/TIFS.2010.2046268

[29] Nwokeji, C. E., Sheikh-Akbari, A., Gorbenko, A., & Mporas, I. (2024). Source camera identification techniques: A survey. *Journal of Imaging*, *10*(2), 31. https://doi.org/10.3390/jimaging10020031

[30] Bennabhaktula, G. S., Timmerman, D., Alegre, E., & Azzopardi, G. (2022). Source camera device identification from videos. *SN Computer Science*, *3*(4), 316. https://doi.org/10.1007/s42979-022-01202-0

[31] Mandelli, S., Bonettini, N., & Bestagini, P. (2022). Source camera model identification. In H. T. Sencar, L. Verdoliva, & N. Memon (Eds.), *Multimedia forensics* (pp. 133–173). Springer, Singapore. https://doi.org/10.1007/978-981-16-7621-5_7

[32] Liu, Y. Y., Chen, C., Lin, H. W., & Li, Z. (2024). A new camera model identification method based on color correction features. *Multimedia Tools and Applications*, *83*(10), 29179–29195. https://doi.org/10.1007/s11042-023-16693-z

[33] Lu, J., Li, C., Huang, X., Cui, C., & Emam, M. (2024). Source camera identification algorithm based on multi-scale feature fusion. *Computers, Materials & Continua*, *80*(2), 3047–3065. https://doi.org/10.32604/cmc.2024.053680

[34] Sameer, V. U., Mukhopadhyay, S., Naskar, R., & Dali, I. (2023). Source camera identification and detection in digital videos through blind forensics. *arXiv Preprint: 2309.03353*.

[35] Song, L., & Huang, H. (2023). Fixed pattern noise removal based on a semi-calibration method. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *45*(10), 11842–11855. https://doi.org/10.1109/TPAMI.2023.3274826

[36] Ferrara, P., Iuliani, M., & Piva, A. (2022). PRNU-based video source attribution: Which frames are you using? *Journal of Imaging*, *8*(3), 57. https://doi.org/10.3390/jimaging8030057

[37] Roy, P., Mitra, S., & Das, N. (2022). Source camera identification using GGD and normalized DCT model-based feature extraction.

In N. Das, J. Binong, O. Krejcar, & D. Bhattacharjee (Eds.), *Proceedings of international conference on data, electronics and computing*(pp. 289–301). Springer Singapore. https://doi.org/10.1007/978-981-99-1509-5_27

[38] Lukas, J., Fridrich, J., & Goljan, M. (2006). Digital camera identification from sensor pattern noise. *IEEE Transactions on Information Forensics and Security*, *1*(2), 205–214. https://doi.org/10.1109/TIFS.2006.873602

[39] Chen, M., Fridrich, J., Goljan, M., & Lukas, J. (2008). Determining image origin and integrity using sensor noise. *IEEE Transactions on Information Forensics and Security*, *3*(1), 74–90. https://doi.org/10.1109/TIFS.2007.916285

[40] Ruderman, D. L. (1994). The statistics of natural images. *Network: Computation in Neural Systems*, *5*(4), 517–548. https://doi.org/10.1088/0954-898X_5_4_006

[41] Tan, X., Lai, S., Liu, Y., & Zhang, M. (2014). Green channel guiding denoising on Bayer image. *The Scientific World Journal*, *2014*(1), 979081. https://doi.org/10.1155/2014/979081

[42] Fridrich, J. (2013). Sensor defects in digital image forensic. In H. T. Sencar & N. Memon (Eds.), *Digital image forensics: There is more to a picture than meets the eye* (pp. 179–218). Springer New York. https://doi.org/10.1007/978-1-4614-0757-7_6

[43] Gloe, T., & Böhme, R. (2010). The 'Dresden image database' for benchmarking digital image forensics. In *Proceedings of the 2010 ACM Symposium on Applied Computing*, 1584–1590. https://doi.org/10.1145/1774088.1774427

[44] Liu, C., Cheng, S., Ding, W., & Arcucci, R. (2025). Spectral cross-domain neural network with soft-adaptive threshold spectral enhancement. *IEEE Transactions on Neural Networks and Learning Systems*, *36*(1), 692–703. https://doi.org/10.1109/TNNLS.2023.3332217

How to Cite: Roy, P., Mitra, S., & Das, N. (2025). Robust Source Camera Identification of Image Using Mean Subtracted Contrast Normalization for Digital Forensic Investigation. *Artificial Intelligence and Applications*. https://doi.org/10.47852/bonviewAIA52023454