Enhancing Source Camera Identification through Higher-order Wavelet Statistics

Dulangi Anuththara Gamage[1], Kasun De Zoysa[2], Asanka Sayakkara[3]
University of Colombo School of Computing, Sri Lanka
[1]anuththaragamage104@gmail.com [2]kasun@ucsc.cmb.ac.lk [3]asa@ucsc.cmb.ac.lk

***Abstract*—In digital forensics, source camera identification (SCI) is an emergent problem that focuses on determining the camera that has been used to capture a given image. Unique characteristics of cameras, such as photo response non-uniformity (PRNU) noise, has been demonstrated to be useful in distinguishing between very similar cameras. Building on existing work, this study introduces a method to uniquely identify source cameras by using statistical features of PRNU noise embedded in images. Here, the PRNU noise is estimated by taking the difference between the camera output and a denoised image. Afterwards, higher order wavelet statistics extraction (HOWS) features are extracted to identify statistical relations between the images taken from the same camera. The proposed method is evaluated under 3 scenarios on the data collected from cameras of smartphones. When distinguishing between cameras from different brands and models, the proposed method produces an accuracy of 95%. In the case of identifying between cameras of different models of the same brand, an accuracy of 92.5% was achieved. In the extreme case of distinguishing between cameras of the same make and model, an accuracy of 85% was achieved. The results also indicate that the proposed method is robust against basic image manipulations.**

***Index Terms*—image forensics, source camera identification, higher order wavelet statistics, photo response non-uniformity noise**

I. Introduction

A concern that arises with the advancement of digital image acquisition and enhancement is the growth of felonies related to pictures and whether they can be trusted enough in the courtroom to be presented as crucial evidence. When it comes to crimes such as child pornography, scientific fraud, insurance claims, and movie piracy cases, it is vital to have reliable methods to establish the digital images' origin and integrity. But the tasks of authentication and validation of images have become challenging since there can be various possible origins to those digital pieces and they might have been subjected to multiple alterations [1], [2]. Due to this reason, one of the most striking problems in the field of digital image forensics is, given an image, identifying the exact device, i.e., source camera, that has been used to capture it. In situations involving child pornography, for example, source camera identification can be used to determine if the perpetrator created or simply owned the incriminating image.

In order to address this problem, it is necessary to identify and analyse features that reveals the identity of a camera device. These features can be considered as two interconnected characteristics: (1) discriminating the devices on model and brand based on the properties that are common to all of them, i.e., camera model identification, and (2) identifying individual devices apart from others in the same model and brand, i.e., individual camera identification [3]. A lot of research efforts have already been put on designing techniques to identify make and model of image-acquisition devices.In contrast, the research presented in this paper focuses on developing a method that can identify the source device of an image even if the cameras under suspicion are of the same model and make, i.e., individual camera identification.

Most of the existing methods for source camera identification (SCI) problem utilise passive forensics, where information regarding the camera is unknown.They have used the most intrinsic patterns available in the images, which are created due to deformities in the source camera. Since the complicated imaging pipeline inevitably leaves some unique traces, these methods are feasible for SCI. However, the success rate of these methods depends on the fact that the characteristics used in these methods are unique to each device. It is impossible to build ideal defects-free imaging devices due to the limitations of manufacturing procedures and the properties of materials. Those defects in the sensors used to capture images introduce various noise or disturbance to the image's pixel values. Therefore, if such a pattern of noise or pixels can be obtained and matched with the reference pattern of an imaging device, then the device used to capture an image can be determined.

There are two basic components to pattern noise: fixed pattern noise (FPN) and the photo-response non-uniformity noise (PRNU) [4], [5]. The pixel-to-pixel differences when the sensor array is not exposed to light can be referred to as FPN. Additional intensity being registered due to the thermally generated free charge within a pixel and more dark current created in some of the pixels due to non-homogeneity in material properties introduced during the manufacturing process can be described as the cause of FPN. Hence temperature and exposure directly affect the FPN. But some high-end cameras subtract a dark frame from every image to suppress this noise as FPN is considered an additive noise. Anyhow, in natural images, PRNU can be observed as a dominant component. Pixels having different sensitivities to light due to the non-homogeneity of silicon wafers and imperfections during the sensor manufacturing process or pixel non-uniformity is what causes PRNU in images. Therefore PRNU is not dependent on humidity or ambient temperature, unlike FPN. Due to those properties and the origin of the PRNU noise, it is improbable that sensors made by the same manufacturer, i.e., same wafer, would reflect correlated PRNU patterns.

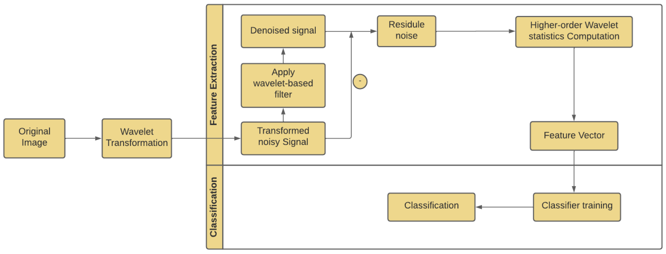

Fig. 1: The overall workflow of the proposed method to distinguish source camera.

II. Related work

In the work of [6], twelve cameras of the *Trust* brand were examined, and errors that are visible in charge-coupled device (CCD) were used to determine the possible source camera of the images. As these defects in the pixels are randomly located in the CCD, that pattern was used as a reference in determining the camera. However, they also have stated that these noises and pixel defects present in the images can be moved or suppressed by image compression algorithms. Moreover, it is crucial to identify the exact cause of the pixel defects as such defects can be shared among cameras from the same batch. However, the problem with this technique is that the defects in the pixels could be absent or cannot be seen in the images under the study and different locations after a lossy compression.

Another artefact that occurs on the images due to the defects in the imaging sensor is noise, which includes FPN and PRNU noise. The dark current noise (FPN), which can only exist if the sensor is not exposed to light, limits the method that it can be utilised for camera identification as it is not possible to use FPN-based methods on regular (non-dark) frames. Nevertheless, in [7], they amplify the FPN by accumulating a large number of frames to conceal the effect of the random noise and use it to recognise the camcorder from videotaped images.

In the work by [4], an approximate PRNU noise is initially calculated and used as the camera reference pattern of the suspected camera. In general, extracting the PRNU noise using processes like flat fielding is not possible. Hence, they have stated that by averaging multiple images, they can approximate the PRNU noise. Then, by calculating the correlation between the camera reference pattern and noise residual of the query image, the presence of the reference pattern in the image is established. In order to reduce the false rejection rate (FRR), the Neyman-Pearson approach [4] is used to calculate a threshold while setting a boundary on the false acceptance rate (FAR). Finally, the correlation value is compared to the threshold. They have also found that the average value of correlations between the reference pattern and image noise residual is reduced due to the JPEG compression.

In [8], a special attention has been given to mobile phone cameras. In their approach, as the fingerprint of the sensor, they have considered each noise pattern using a bio-metric analogy. They also have classified the images captured from mobile phones using sensor pattern noise (SPN). They have used a wavelet-based feature vector to characterise the fingerprint. For the classification purpose, 81 features that characterise the sensor fingerprint are extracted. Then classification was performed using a support vector nachine (SVM) with RBF kernel just as in their previous work.

In [9], a technique was presented that can be used to deal with the scene interference and enhance the SPN effectively. They proposed five models that can be used with the wavelet-based denoising filter and applied to the extracted unenhanced SPN. Despite introducing these five models, they do not provide any guideline for selecting the enhancing model as there is currently no theory for modelling SPN and scene details, hence no theoretical basis for selecting the optimal model.

In the beginning of source camera identification, [10] demonstrated a method to identify images acquired by a small number of camera models. The colour features they proposed for SCI are red-green-blue (RGB) pairs correlation, average pixel value, RGB pairs energy ratio, neighbour distribution centre of mass, and wavelet domain statistics. In addition to those colour features, quality features of images in image quality metrics (IQM) were also used to feed to the classifier to aid in distinguishing between cameras.

The approach proposed by [11] for camera source identification works with high performance by using wavelet features and bi-coherence as input to the classifier. To depict the distinctive distortions that different cameras cause in their images, they have suggested an identification approach that uses bi-coherence magnitude, phase statistics and wavelet coefficient statistics. Here, the impact of non-linear distortions generated by the imaging pipeline on higher-order picture statistics and the impact of image processing procedures on the wavelet domain is considered. They use the sequential forward feature selection approach to lessen the correlation between characteristics and increase the accuracy of source camera identification after extracting statistical features of image bi-coherence and wavelet coefficient features. Then, an SVM was utilised to feed the above features. Following that [12] proposed another method in which they extract wavelet coefficient co-occurrence features and higher-order wavelet features from captured images. Then, to minimise the correlation and redundancy of the features, the sequential forward feature selection (SFFS) algorithm is used, and classification is performed using a multi-class SVM to identify the source of the image. The effectiveness of the suggested approach is empirically improved on photos, even when compared to previous ways.

In [13], the authors have combined the forensic features such as binary similarity measures, image quality features, and higher-order wavelet statistics to determine the camera of the image under investigation. With the three sets of features, they have experimented with two ways: (1) merge all of these contrasting features and build one classifier (feature fusion), and (2) handle the characteristics separately with appropriate classifiers before fusing them at the decision level(decision level fusion).

Original pictures, residual noise images, and residual noise contourlet transform coefficients are all used to recover local binary pattern (LBP) features in [14] respectively. Then, the local phase quantisation (LPQ) features are retrieved from the original and residual noise pictures. The combined LBP and LPQ features are then sent to the multi-class lib-SVM classifier after both LBP and LPQ features are extracted from H and V colour channels in hue-saturation-value (HSV) colour space.

TABLE I: Cameras used in the experiment and their properties

Camera Model	Resolution	Image Format	No. of images
Samsung Galaxy S9	4032x3024	JPEG	300
Samsung Galaxy S9	4032x3024	JPEG	300
Galaxy A50	1860x4032	JPEG	300
Galaxy J7 Max	4128x2322	JPEG	300
Huawei Y7	3000x4000	JPEG	300

III. Methodology

The majority of source camera identification techniques based on PRNU noise use the correlation between the camera fingerprint and the PRNU noise pattern of the particular image. However, correlation-based approaches are highly influenced by random noise components and the scene details in the estimated PRNU. When photos undergo minor visual and geometric alterations, the identification process requires more work. In order to perform correlation between distinct photos, it is also necessary to use photographs of equal size. Furthermore, traditional feature-based techniques necessitate a high number of features to obtain good recognition accuracy, requiring the extraction of more discriminating characteristics from the images. To solve these issues, feature-based methods that use sensor pattern noise introduce a robust source camera identification model. The Figure 1 illustrates the overall workflow of the proposed method.

A. Preparation of the Dataset

There are various publicly available datasets for image processing, such as Dresden and VISION. However, in order to cover modern-day cameras with the latest technological advances, it was decided to prepare a new dataset for the purpose of this research. The smartphone cameras illustrated in the Table I were used for creating the dataset and performing experiments accordingly.

B. Estimating Noise

In order to distinguish the exact source camera of an image among many potential cameras, it is necessary to identify and extract the attributes that are unique to each camera. Each image is susceptible to many sorts of noise, some of which are unique to the equipment that captured it. Among them, pattern noise — more specifically PRNU — is a deterministic noise component in photographs that is consistent across all photos taken by the particular camera. Moreover, since different sensors have their own PRNU patterns and no two sensors would have similar PRNU patterns, it is more suitable for the identification of the source camera.

The Equation 1 illustrates the mathematical model of acquiring an image as described by [4].

$$o_{ij} = f_{ij}(p_{ij} + s_{ij}) + d_{ij} + a_{ij} \tag{1}$$

In Equation 1, o_{ij} is the sensor's output, f_{ij} is our interest, the PRNU noise component, p_{ij} is the number of photons that hit the sensor, a_{ij} is the additive random noise, d_{ij} is the dark current, and s_{ij} is the shot noise. Unfortunately, extracting the SPN straightly from the real output o_{ij} is challenging. Instead, it is estimate the noiseless output by de-noising the real sensor output p_{ij}. Then denoising filters are used to denoise the photos. To obtain the prominent component PRNU or the noise residual, denoised image $I_{denoised}$ is subtracted from the original image I.

$$I_{noise} = I - I_{denoised} \tag{2}$$

The quality of the extracted PRNU directly impacts the final detection accuracy when determining the source camera. However, the estimated PRNU, on the other hand, is frequently corrupted by random noise and the contents of the picture, lowering detection accuracy. It is critical to use a proper filter for denoising in order to obtain a pure PRNU. The benefit of utilizing the noise residual is that any low-frequency components in the pictures that are not sensor-specific are automatically muted. Furthermore, higher discriminating capability can be seen by the statistical features of the estimated PRNU in comparison to the statistical characteristics of the original image.

C. Higher order wavelet statistics extraction

HOWS characteristics are used in the proposed feature-based technique for source camera identification, allowing the coefficient distribution of predicted PRNU noise to be captured. These features are capable of detecting regularities in images and have been proven to be beneficial in a variety of forensic applications. HOWS is a statistical model based on multiscale wavelet decomposition that is used to analyze natural photographs. This model considers the first and higher-order statistics to capture the regularities in the images.

The decomposition of PRNU images using separable Quadrature Mirror Filters (QMF), which divide the frequency space into different scales and orientations, is the subject of this study. Decomposition is applied to each channel separately in colour photographs. For each sub-band coefficient, statistical properties such as kurtosis, skewness, variance and mean must be determined.

D. Classification

As per current findings, support vector machines, which are extensively used as supervised learning methods for classifying two or more sets of data, were used for classification. SVM can minimise classification error while simultaneously increasing the separation between classes (geometric margin). When only test data attributes are presented to the system, SVM develops a model based on the training data that can predict the target value for the test data. When the data is not linearly separable, the kernel functions are used to map the features into a higher-dimensional space, making the data linearly separable. The classifier's performance can be evaluated with different cost parameters and kernel functions to choose the optimal parameters.

(a) Denoised image using Mihcak Filter (b) Noise residual extracted from image (c) Noise residual extracted from clear blue sky

Fig. 2: Images representing noise residual extraction.

TABLE II: Configuration of experimental setup

System Model	MSI GF63 Thin 9SC
Processor	Intel(R) Core(TM) i7-9750H CPU @ 2.60GHz
OS Name	Microsoft Windows 10 Home Single Language
System Type	x64-based PC
Memory	8GB
MATLAB version	R2021a

IV. Result & Evaluation

This research is developed using the MATLAB R2021b trial version provided by MathWorks software corporation[15]. MATLAB product provides the necessary apps and functions for all the research requirements as it is specialized in mathematical computing. Along with the MATLAB, other Mathworks products such as Image Processing Toolbox 11.4, Statistics and Machine Learning Toolbox 12.2 and Wavelet Toolbox 6.0 products are also incorporated to implement the functionalities in this work. Moreover, for several other experimental purposes, the scikit-learning library is also used along with the Google Colaboratory. Moreover, for the purpose of classification, LIBSVM, which is a library for support vector machines, is also used. The personal computer with the configurations mentioned in the Table II is used for the experiments carried out, which provides the necessary processing power.

Dataset was prepared using several smartphone brands and models. Brands include Samsung and Huawei. Detailed descriptions of the images taken from those cameras are mentioned in the Table I. Two smartphones that have the same make and model have been used to analyze and evaluate the performance of the model in that particular case of investigation. Due to the concerns regarding the quality of the image, image compression in various image saving formats and popularity, the JPEG format of all the images is utilized. The images used for training are routine shots from natural scenarios which do not contain many scene details.

A. PRNU Estimation

Several denoising filters have been tried to extract the noise residuals from the images, including a Gaussian-based denoising filter and various other spatial domain and wavelet domain filters. Due to the ability to suppress the scene contents and most of the pattern noise, wavelet domain denoising filter, Mihack's filter [16] is finally chosen to extract the PRNU noise. A denoised image and the extracted noise residuals using Mihack's filter can be seen in the Figure 2. Moreover, the wavelet coefficients of each individual colour channel of the noise residual are also computed.

B. HOWS features Extraction

The noise residuals of all the images in the data set are used to compute HOWS features. The db4 wavelet is used to conduct three-level wavelet decomposition on each colour channel of the image individually, resulting in nine subbands. For all subbands, first and higher-order statistical features such as kurtosis, skewness, variance and mean are computed, totalling 108 features. These characteristics clearly distinguish between photos captured by various cameras, allowing them to be classified according to their source category.

C. Classification

LIBSVM, a software developed mainly to support vector classification, is integrated to carry out the classification. It supports different SVM formulations and also provides multi-class classification. LIBSVM tools provide interfaces and extensions to most programming languages, including an interface to MATLAB as well. The LIBSVM version 3.25 is used in this research work to implement a classifier. In this work, two-class SVM is used to classify two-class SCI problems, and multi-class SVM is used to evaluate the model's performance with an increased number of source cameras.

In this study, the nine-tenth of the dataset is used as the training data, and the remaining one folds are used as the testing data to assess the classifier. The number of class samples picked from two different classes is nearly equal in

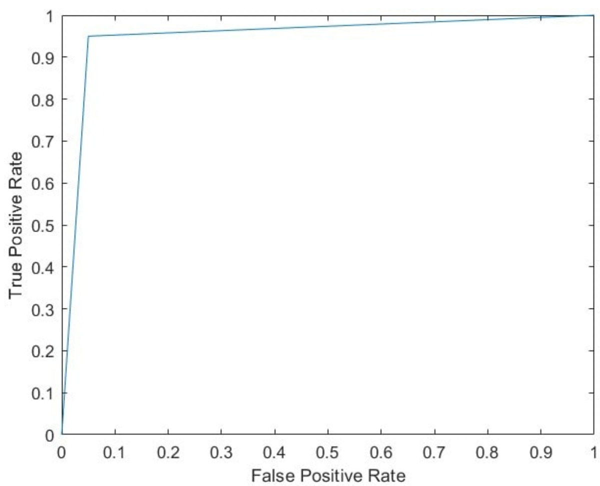

Fig. 3: ROC curve of the scenario of different camera brand and model.

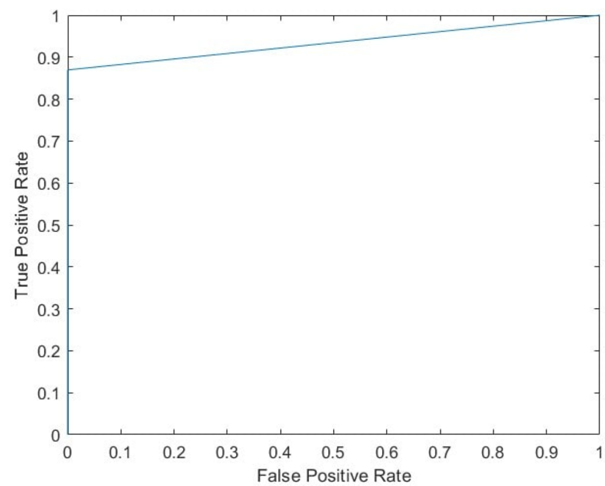

Fig. 4: ROC curve of the scenario of same camera brand and different models.

each fold. The performance of the model is measured using the identified accuracy and visualized using the confusion matrix and ROC curves.

D. Analysis of detection

Evaluation of the source camera identification can be carried out in three categories to determine whether the proposed method satisfactorily serves the problem. As the difficulty in detection gradually increases with the camera brand, model and individual device, such an experiment model is suited.

1) Differentiating camera models and brands - Samsung S9 and Huawei Y7
2) Differentiating camera models from same brand cameras - Samsung S9 and Samsung J7 Max
3) Differentiating individual cameras from same make and model - Samsung S9

TABLE III: Confusion table of the scenario of different camera brand and model.

	Actually Positive	Actually Negative
Predicted Positive	19	1
Predicted Negative	1	19

TABLE IV: Confusion table of the scenario of same camera brand and different models.

	Actually Positive	Actually Negative
Predicted Positive	17	0
Predicted Negative	3	20

TABLE V: Confusion table of the scenario of same camera model and brand.

	Actually Positive	Actually Negative
Predicted Positive	15	5
Predicted Negative	1	19

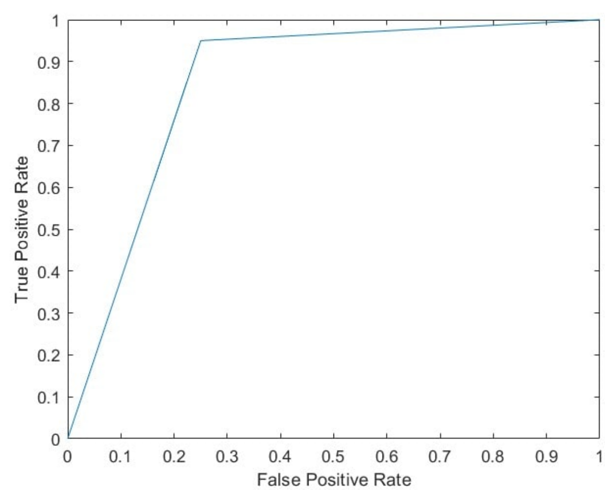

Fig. 5: ROC curve of the scenario of same camera model and brand.

TABLE VI: Performance evaluation of different cases

No.of classes	Camera models	Average accuracy
2	Samsung S9,Huawei Y7	95%
2	Samsung S9,Samsung J7	92.5%
2	Samsung S9,Samsung S9	85%

Table VI shows the results obtained for the classification accuracy. For case 1, case 2 and case 3, the average accuracy of 95%, 92.5%, and 85% is obtained and respective confusion matrices can be seen in Tables III, V, and IV. ROC curves in Figures 4, 5, and 3 indicate how well the models distinguish classes. These findings indicate that the proposed source camera identification algorithm can identify the source camera regardless of the camera's manufacturer or model. When two different make/models of source cameras are utilized in the experiment, the performance improves. However, when two similar cameras are used, the performance decreases.

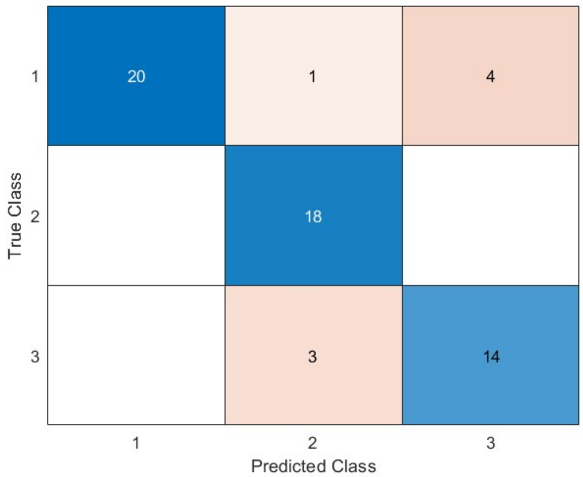

Fig. 6: Confusion Matrix for Multi Class Classification

E. Performance for multi-class classification

The algorithm's performance is further assessed by escalating the number of cameras utilized in the experiment. The classification is done with a multi-class SVM. When photos from three different cameras were utilized in the experiment, the Figure 6 demonstrates how well the algorithm performed.

It can be shown that this group may reach an average accuracy of 86.66%. When the number of camera classes is increased, the performance suffers slightly. When numerous classes are utilized, there is a tiny overlap of features, which causes this. As a consequence of the results mentioned above, even with a more significant number of camera classes, the features pertaining to distinct camera classes remain separable.

V. Conclusion

This research study identified the features that can be extracted from JPEG images and which of those features can be used to build a fingerprint to the source camera that captures the image. It was found that it is possible to use both the wavelet domain features and the spatial domain for SCI, but spatial domain features are susceptible to various artefacts of the image acquisition process and the scene details of the image. Therefore, higher-order wavelet domain features in various decomposition levels which have the minimum effect from pattern artefacts were used in this work.

For the estimated PRNU noise, HOWS features are generated and used with SVM classifiers to identify the source camera. The results obtained from the experiments carried out using the proposed technique show the accuracy of 95%, 92.4%, and 85% for the cases of investigation of cameras of a different brand, same brand but different model and same brand and exact model, respectively. It is visible from the final results that between cameras from different brands and models, discrimination is considerably high compared to the model's performance when classifying the images taken from the cameras of the same brand and model. This is due to the cameras of the same make and model holding considerable correlation among them, which makes it harder for the model to distinguish them separately. Moreover, since the proposed solution uses statistical features in the wavelet domain instead of a correlation-based approach, the dimensions and locations of the images have the most negligible impact on these statistical measurements. Therefore, it can be concluded that this model's performance would not be affected by alterations like re-sampling and cropping.

This work can be further improved to discriminate between the cameras from the same brand and model. In order to do that, wavelet domain features can be used along with the spatial domain features as well. The proposed model also can be improved to perform better even with the multiple cameras under investigation. It will also be a significant contribution if the proposed model can detect the forged PRNU patterns in an image.

References

[1] Z. B. Parry, "Digital manipulation and photographic evidence: Defrauding the courts one thousand words at a time," 2009.

[2] A. Rocha, W. Scheirer, T. Boult, and S. Goldenstein, "Vision of the unseen: Current trends and challenges in digital image and video forensics," *ACM Comput. Surv.*, vol. 43, no. 4, oct 2011. [Online]. Available: https://doi.org/10.1145/1978802.1978805

[3] B. Xu, X. Wang, X. Zhou, J. Xi, and S. Wang, "Source camera identification from image texture features," *Neurocomputing*, vol. 207, pp. 131–140, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231216303319

[4] J. Lukas, J. Fridrich, and M. Goljan, "Digital camera identification from sensor pattern noise," *IEEE Transactions on Information Forensics and Security*, vol. 1, no. 2, pp. 205–214, 2006.

[5] T. Filler, J. Fridrich, and M. Goljan, "Using sensor pattern noise for camera model identification," in *2008 15th IEEE International Conference on Image Processing*, 2008, pp. 1296–1299.

[6] Z. J. Geradts, J. Bijhold, M. Kieft, K. Kurosawa, K. Kuroki, and N. Saitoh, "Methods for identification of images acquired with digital cameras," in *Enabling Technologies for Law Enforcement and Security*, S. K. Bramble, E. M. Carapezza, L. I. Rudin, L. I. Rudin, and S. K. Bramble, Eds., vol. 4232, International Society for Optics and Photonics. SPIE, 2001, pp. 505 – 512. [Online]. Available: https://doi.org/10.1117/12.417569

[7] K. Kurosawa, K. Kuroki, and N. Saitoh, "Ccd fingerprint method-identification of a video camera from videotaped images," in *Proceedings 1999 International Conference on Image Processing (Cat. 99CH36348)*, vol. 3, 1999, pp. 537–540 vol.3.

[8] A. Sandoval Orozco, J. Hernandez-Castro, L. García Villalba, S. Gibson, D. González, and J. Rosales, "Smartphone image acquisition forensics using sensor fingerprint," 06 2015.

[9] C.-T. Li, "Source camera identification using enhanced sensor pattern noise," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 2, pp. 280–287, 2010.

[10] M. Kharrazi, H. Sencar, and N. Memon, "Blind source camera identification," in *2004 International Conference on Image Processing, 2004. ICIP '04.*, vol. 1, 2004, pp. 709–712 Vol. 1.

[11] F. Meng, X. Kong, and X. You, "A new feature-based method for source camera identification," in *IFIP International Conference on Digital Forensics*. Springer, 2008, pp. 207–218.

[12] B. Wang, Y. Guo, X. Kong, and F. Meng, "Source camera identification forensics based on wavelet features," 09 2009, pp. 702–705.

[13] O. Celiktutan, B. Sankur, and I. Avcibas, "Blind identification of source cell-phone model," *Trans. Info. For. Sec.*, vol. 3, no. 3, p. 553–566, sep 2008. [Online]. Available: https://doi.org/10.1109/TIFS.2008.926993

[14] B. Wang, Y. Guo, X. Kong, and F. Meng, "Source camera identification forensics based on wavelet features," in *2009 Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, 2009, pp. 702–705.

[15] "Matlab." [Online]. Available: https://in.mathworks.com/products/matlab.html

[16] M. Mihcak, I. Kozintsev, and K. Ramchandran, "Spatially adaptive statistical modeling of wavelet image coefficients and its application to denoising," vol. 6, 04 1999, pp. 3253 – 3256 vol.6.