IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

SPECIAL SECTION ON DIGITAL FORENSICS THROUGH MULTIMEDIA SOURCE INFERENCE

Received June 30, 2019, accepted July 4, 2019, date of publication July 11, 2019, date of current version August 9, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2928356

Daxing Smartphone Identification Dataset

HUAWEI TIAN[1], (Member, IEEE), YANHUI XIAO[1], (Member, IEEE), GANG CAO[2], (Member, IEEE), YONGSHENG ZHANG[1], ZHIYIN XU[1], AND YAO ZHAO[3], (Senior Member, IEEE)

[1]Research Center of Public Security Information, People's Public Security University of China, Beijing 100038, China
[2]School of Computer Science and Cybersecurity, Communication University of China, Beijing 100024, China
[3]Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China

Corresponding author: Yanhui Xiao (xiaoyanhui@ppsuc.edu.cn)

This work was supported in part by the National Science Foundation of China under Grant 61772539, Grant 6187212, and Grant 61401408, in part by the Technical Research Program of the Ministry of Public Security of China under Grant 2017JSYJC01, in part by the Science and Technology Projects of Sichuan Province under Grant 2018JY0521, and in part by the 2019 First-Class Subjects of PPSUC under Gant 2019XK0107.

ABSTRACT Over the past few years, the imaging device has changed from digital cameras to smartphone cameras. With the popularity of mobile Internet applications, there explode massive digital images and videos captured by such smartphones, which are nearly held one per person. Consequently, the capturing source of images/videos delivers valuable identity information for criminal investigations and critical forensic evidence. It is significant to address the source identification of smartphone images/videos. In this paper, we build a Daxing smartphone identification dataset, which collects images and videos from extensive smartphones of different brands, models and devices. Specifically, the dataset includes 43 400 images and 1,400 videos captured by 90 smartphones of 22 models belonging to 5 brands. For example, there are 23 smartphone devices for the iPhone 6S (Plus) model. To the best of our knowledge, Daxing dataset uses the largest amount of smartphones for image/video source identification compared with other related datasets, as well as the highest numbers of devices per model and captured images/videos. The dataset has been released as a free and open-source for scientific researchers and criminal investigators.

INDEX TERMS Image forensics, video forensics, source identification, benchmarking.

I. INTRODUCTION

In the past decade, smartphones have gained popularity due to advantages of economy, portability and communication convenience. The most widely used imaging device has changed from digital compact and single lens reflect cameras to smartphone cameras. Since the smartphone almost cannot afford to get away for a moment, it can provide some key information for criminal investigation and critical forensic evidence. The blind source identification or inference from output data of smartphones, i.e., images, videos and audios, has been a hot topic in digital forensics research field. Building a standard testing dataset of smartphone-captured images/videos is essential and significant for promoting such a research.

The first dataset adopted for image forensics research is UCID (Uncompressed Color Image Dataset) [1], which consists of 1,338 uncompressed TIFF images taken by one Minolta Dimage 5 camera. Dresden dataset [2], [3] is the first public dataset specially built for image source identification. It includes 16,961 JPEG and 1,491 RAW images from 74 cameras of 26 models belonging to 14 brands. There are less than 5 samples for a certain camera brand. RAISE (RAw ImageS datasEt) [4] uses three cameras to capture 156 RAW images from 7 different scenes, i.e. outdoor, indoor, landscape, nature, people, objects and building. VISION (Video and Image dataset for Source Identification) [5] collects 11,732 images and 648 videos taken by 35 smartphones or tablet PCs of 30 models with 11 brands. In the competition of camera model identification (CCMI) held by IEEE Signal Processing Society in 2018 [6], an image dataset including training and test sets is made public. The images in training set are captured by a total of 10 smartphones, 275 images for each smartphone. A dataset is built for HDR image forensics (DHIF) [7], which contains 5,415 HDR images and their corresponding SDR versions from 23 smartphones. The Smartphone Image Denoising Dataset (SIDD) [8] can also be used for image source identification. It contains 30,000 images captured

The associate editor coordinating the review of this manuscript and approving it for publication was Irene Amerini.

by 5 smartphones under 10 illumination scenes. Recently, Galdi *et al.* release a digital image and video database named as SOurce Camera REcognition on Smartphones (SOCRatES), which is specially designed for source camera recognition on smartphones. With 103 different devices, SOCRatES [9] is the source camera identification database that includes the highest number of different sensors. There are also some public databases available for video source identification and content forgery detection. For example, SULFA (Surrey University Library for Forensic Analysis) [10] owns 150 videos from three cameras, each video is about 10 seconds, 30fps. The video forgery detection database (VFDD) [11], [12] is built for video tamper detection, and contains 1495 original videos shoot by 27 devices under 8 scenes. There are also some other datasets built for image/video source identification and tampering detection [13]–[15].

For the research of multimedia source identification, it is very important to study the source forensics technology of camera device identification rather than just camera model identification. The review of existing image/video source identification databases shows that the number of devices of the same model is generally not enough, where SOCRatES reaches the highest, i.e., the number of iPhone 6 is 9, and Dresden reaches 5. In order to attenuate such a deficiency, multiple smartphones of the same model are used in building our Daxing dataset, such as 13 iPhone 6S smartphones and 10 iPhone 6S Plus smartphones. The database contains 43,400 original images and 1,400 original videos, which are captured by 90 cameras of 22 models of 5 brands and all in Daxing District of Beijing, China. To the best of our knowledge, the Daxing dataset includes the largest amounts of smartphones with the same model, and captured images/videos. Daxing dataset is released freely for scientific purposes at https://github.com/xyhcn/Daxing.

The remaining part of the paper is organized as follows. In Section II, we review available datasets for image source identification. Section III provides a complete description of the dataset covering 5 subsections. In Section IV, the dataset is exploited to evaluate source identification applications. Section V draws conclusions.

II. RELATED WORK

In this section, several popular test datasets on the image and video source identification are investigated and analyzed in detail. Their characteristics would be understood deeply.

A. UCID

UCID, as one of the earliest datasets available for image forensics research, consists of 1,338 compressed TIFF images [1]. However, such images are not the unaltered outputs from cameras. They enjoy a rather small resolution of 512×384 or 384×512 pixels, and may suffer some out-camera postprocessing. Moreover, all the images are captured by the same one Minolta Dimage 5 camera.

B. DRESDEN

Dresden dataset is the first major dataset specially built for image source identification [2], [3]. The images are collected by 73 cameras from 25 models of 14 brands. There are generally no more than 5 devices for each model. These cameras include the most popular camera brands on the market, such as Nikon, Canon, Fuji, etc. To avoid damaging the image quality from JPEG compression, all images are stored at the highest quality. Each camera captures multiple images from different angles. The images range in resolution from 3072×2304 to 4352×3264, with a total of 16,961 JPEG images and 1,491 RAW images. The overall quality of images in Dresden database is relatively better, which can provide a benchmark for future digital forensics and references for the development of other databases.

C. RAISE

The RAISE dataset includes 8,156 raw images with a wide variety of semantic contents and technical parameters [4]. Three camera devices (each one for Nikon D40, Nikon D90 and Nikon D7000) yield images at high resolutions (3008×2000, 4288×2848 and 4928×3264). Such images are saved in an uncompressed format (Compress Raw 12-bit and Lossless Compress Raw 14-bit) as primary output of the used cameras. Each image falls into one of the scene categories of outdoor, indoor, landscape, nature, people, objects and buildings [4].

D. VISION

VISION collects images and videos from 35 smartphones or tablet PCs of the 11 brands including Apple, Asus, Huawei, Lenovo, LG electronics, Microsoft, OnePlus, Samsung, Sony, Wiko and Xiaomi [5]. There are 11,732 native images and 648 native videos. In real life, most images and videos are shared via social media platforms, such as Facebook and YouTube. In order to simulate such an application, 7,565 images are compressed by Facebook (including two modes of high quality and low quality) and WhatsApp, resulting in 22,695 compressed images. In addition, 622 videos are compressed using YouTube and 644 videos were compressed using WhatsApp, resulting in 1,266 compressed videos in total.

E. CCMI

IEEE Signal Processing Society organizes a competition on identifying the source of smartphone images [6]. The competition also provided participants with a standard dataset, which was divided into training and test sets. All images in the training set are from 10 smartphones, i.e. Sony NEX-7, Motorola Moto X, Motorola Nexus 6, Motorola DROID MAXX, LG Nexus 5x, Apple iPhone 6, Apple iPhone 4s, HTC One M7, Samsung Galaxy S4, and Samsung Galaxy Note 3. Each smartphone captured 275 images with different scenes, and a total of 2,750 images can be used to train smartphone image source identification algorithm.

In addition, the test set includes 2,640 images from the same 10 smartphones as training set. Half of images in test set have been processed with compression and magnification at different scale, some have been gamma corrected, and all images are cropped to 512×512 pixels.

F. DHIF

Activating the HDR function of smartphones would bring new challenges to the image source identification. The DHIF dataset is constructed for HDR image forensics [7]. It contains 5,415 HDR images and their corresponding SDR collected by 23 smartphones. The image acquisition process consists of three shooting modes of handheld, tripod fixed and shaking, the diversity of image content is also ensured.

G. SIDD

Although the construction of SIDD database mainly focused on the research of smartphone noise elimination, it can obviously be applied to the research of image source forensics because it provides as many as 30,000 original images [8]. SIDD employed 5 smartphones (Apple iPhone 7, Google Pixel, Samsung Galaxy S6 Edge, Motorola Nexus 6, and LG G4) to shoot under 10 different illumination scenes. In addition, the database also provides noise-free images of these native images as ground truth images. Therefore, researchers can use this dataset to benchmark image denoising algorithms.

H. SOCRatES

Different from the previous published databases, SOCRatES [9] is collected by the smartphone owners themselves. As such, great heterogeneity and realness is introduced in the captured data. SOCRatES has about 9,700 images and 1000 videos captured with 103 smartphones of 15 different makes and about 60 different models. With the 103 different devices, SOCRatES has become the source camera identification database including the highest number of different sensors.

I. SULFA

There are few video datasets designed for source forensics, and SULFA is the one created by the University of Surrey [10]. It collects 150 videos, each 10-s long, at 30 fps with a resolution of 320 × 240 pixels. The native videos are given compressed in H.264/AVC and MJPEG, for each camera, namely, a Canon SX220, a Nikon S3000, and a Fujifilm S2800HD. Authors built the dataset to support the research on cloning detection, performed by means of Adobe Photoshop CS3 and Adobe After Effect CS5.

J. VFDD

VFDD is created for video tamper detection. It contains 1495 original videos shot by 27 devices in 8 different scenes [11], [12]. The VFDD (version 1.0) released in 2017 containing 505 original videos shot by 12 devices in 8 different scenes. After editing them, 135 tampered videos were obtained, all these added up to 640 videos. In the VFDD (version 2.0) released in 2018, 15 new devices (including smartphones and cameras) were added. 990 original videos were shot in 8 different scenes, and 262 videos were selected for editing, resulting in 560 tampered videos. All these added up to 1,550 videos.

K. DAXING

In general, it is more important for completing the task of individual-level device identification than that of model-level identification. To achieve that end, a large number of different smartphones with the same model are required to create the smartphone identification dataset. However, the above review reveals that the number of devices with the same model is typically not enough in the existing source identification databases. In order to attenuate such a deficiency, many smartphones with the same model are adopted in building our Daxing dataset, such as 13 iPhone 6S smartphones and 10 iPhone 6S Plus smartphones. Furthermore, to the best of our knowledge, the quantities of captured images and videos in Daxing dataset achieve the highest, which are 43,400 for original images and 1,400 for original videos, respectively.

III. DATASET DESCRIPTION

The creation of image database mainly includes the following four phases: 1) the selection of smartphone brand and model; 2) image and video capturing, including method and content; 3) image and video storage and encoding; 4) application guidance of Daxing dataset.

A. SELECTION OF SMARTPHONE BRAND AND MODEL

We selected the popular smartphone brands on the market currently for image collection, including Huawei, Apple, OPPO, VIVO, and Xiaomi. Each model has at least 5 smartphone devices. The details are shown in Table 1.

B. IMAGE COLLECTION

After the devices are selected, the image collection work needs to be carried out. The image collection work mainly includes the selection of the shooting scene, the photographing setting of the imaging device, and the number of images. Selected scenes include ''sky'', ''grass'', ''stone'', ''trees'', ''staircase'', ''indoor vertical printer'', ''lobby wall'', ''white wall in the classroom'', etc. In each scene, images are captured when smartphone is placed at three different angles (we regard the vertical position of the smartphone as the reference, and the placement angle is set as 90 degrees, 0 degrees when rotating 90 degrees counterclockwise and 180 degrees when rotating 90 degrees clockwise, respectively). All of cameras are set to ''Default'' mode and flash was set to ''Off'' mode in the collection process. Specific information including scenes display angles and image numbers is shown in Table 2. The ''sky'' scene provides the most amount of images, no less than 102 per smartphone. For the other seven scenes, there are no less than 50 photos of each scene. As a

TABLE 1. Brand, model and number of the smartphones used in creating Daxing dataset.

Brand	Model	Number
Huawei	Huawei P20	5
	Huawei Mate 9	3
	Huawei Mate 9 Plus	1
	Huawei P9	5
	Huawei P10	3
	Huawei P10 Plus	6
Apple	iPhone 6	5
	iPhone 6 Plus	4
	iPhone 6S	13
	iPhone 6S Plus	10
	iPhone 7	3
	iPhone 7 Plus	5
	iPhone 8 Plus	2
OPPO	OPPO R9	6
	OPPO R9S PLUS	1
	OPPO R11	5
	OPPO R11T	1
VIVO	VIVO X9	4
	VIVO X9 Plus	1
	VIVO X9I	1
	VIVO Y85	1
Xiaomi	Xiaomi 4A	5

TABLE 2. Number of images captured with different Scenes and camera settings.

Scene	Angle	NO.	Mode	Flash light	Total amount
sky	0	≥26	Default	Off	≥102
	90	≥50	Default	Off	
	180	≥26	Default	Off	
grass, stone, trees, staircase, indoor vertical printer	0	≥13	Default	Off	≥5×56
	90	≥30	Default	Off	
	180	≥13	Default	Off	
lobby wall, white wall in the classroom	0	≥13	Default	Off	≥2×50
	90	≥24	Default	Off	
	180	≥13	Default	Off	

result, each smartphone captures no less than 482 images in 8 scenes. Figure 1 shows some examples in the Daxing dataset.

C. IMAGE STORAGE AND CODING

All of images in Daxing dataset are directly copy from smartphones without postprocessing. The resolution of the images is the default resolution of each smartphone camera. The resolution of smartphone is shown in Table 3. For the convenience, we have encoded all devices. In the encoding process, the first code of all smartphones of the same brand is identical, the first two codes of all smartphones of the same model are identical, and the last two codes represent different individual devices of the same model. The specific codes are shown in Table 3.

D. VIDEO COLLECTION, STORAGE, AND CODING

The video collection work mainly includes the selection of the shooting scene, the photographing setting of smartphone, and the number of captured videos. The selected scenes typically include ''sky'', ''grass'', ''stone'', ''trees'', ''staircase'', ''indoor vertical printer'', ''lobby wall'', ''white wall in the classroom'', etc. In each scene, videos are captured when smartphone is placed vertically. At least three videos are captured in each scene. Moreover, all videos take longer than 10 seconds to shoot. All cameras are set to ''Default'' mode. All videos in this dataset are directly copied from smartphones without postprocessing. The resolution of the video adopts the default setting of each smartphone camera, as listed in Table 3. For the convenience, we have encoded each device. In the encoding process, the first code of all smartphones of the same brand is identical, the first two codes of all smartphones of the same model are identical, and the last two codes represent different individual devices of the same model. The specific codes are also shown in Table 3.

E. APPLICATION GUIDANCE OF DAXING

Daxing is suitable for fingerprint extraction, image source identification, and video source identification. Firstly, in order to satisfy fingerprint extraction, each smartphone captures no less than 482 images in 8 scenes. Specifically, the ''sky'' scene provides the most amount of images, no less than 102 per smartphone. The reason behind attributes that the fingerprint extracted from flat blue sky images owns higher quality. Furthermore, the smartphone may be placed horizontally and vertically in shooting images. In order to extract fingerprints from multiple images conveniently, we create three folders for each smartphone, which are named according to the angle of smartphone placing.

Secondly, in order to satisfy image and video source identification, the first code of all smartphones of the same brand is identical. The first two codes of all smartphones of the same model are identical, and the last two codes represent different individual devices of the same model. In this way, researchers can conveniently use the dataset for brand-level, model-level and individual-level image and video source identification.

IV. EXPERIMENTAL EVALUATIONS

The development of the database provides researchers who are engaged in image and video source forensics with a standard test dataset. In this section, the images and videos collected will be used to test device fingerprint extraction and

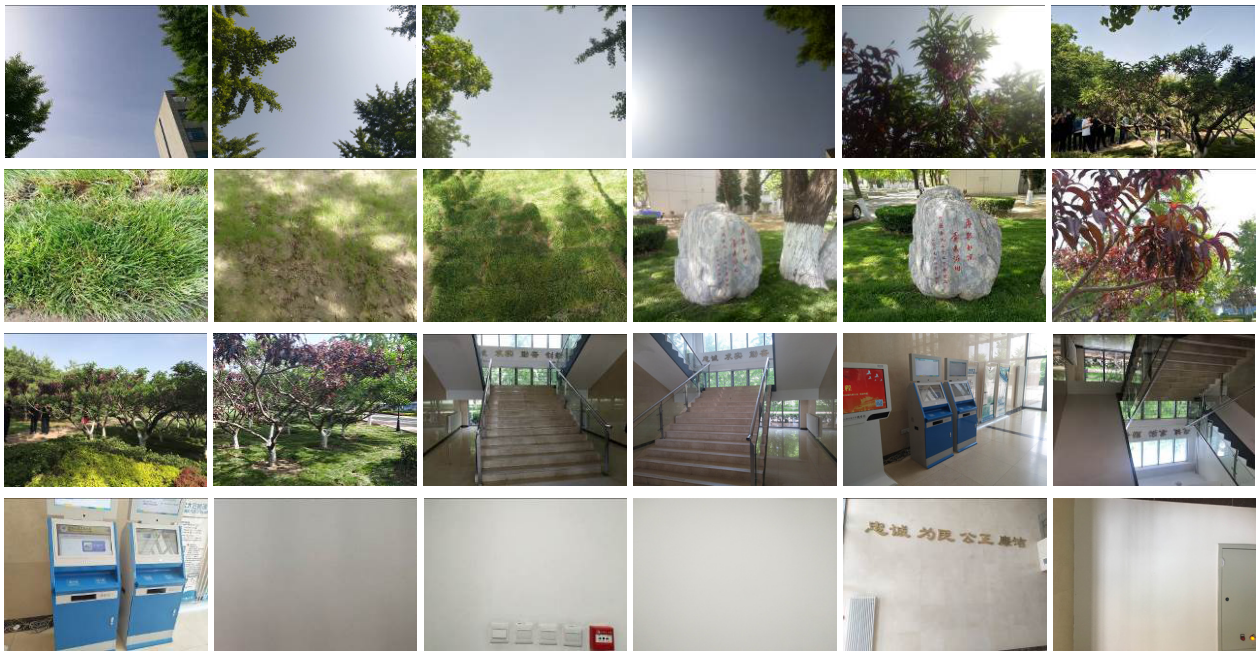


FIGURE 1. Some examples in Daxing dataset.

source identification technology based on Photo-Response Non-Uniformity noise (PRNU).

A. FINGERPRINT EXTRACTION AND SOURCE IDENTIFICATION

PRNU noise is caused by imperfections during sensor manufacturing process and inhomogeneity of silicon wafers. PRNU noise $\mathbf{K} \in \mathbb{R}^{w \times h}$ caused by imperfections of sensor is very weak, its resolution $w \times h$ is same as resolution of the sensor. The output of the sensor $\mathbf{I}$ can be expressed as follows [16]:

$$\mathbf{I} = \mathbf{I}_0 + \mathbf{K}\mathbf{I}_0 + \Theta \tag{1}$$

where $\mathbf{I}_0$ is the original input image, $\mathbf{I}$ is the output image, and Θ is random noise. Here, PRNU noise $\mathbf{K}$, which is a multiplicative noise, operates on $\mathbf{I}_0$, and its distribution is similar to AWGN. The rich frequencies and content of PRNU noise $\mathbf{K}$, and the uniqueness of sensor can be view as a device fingerprint for image source camera identification and image forgery detection.

Device fingerprint $\mathbf{K}$ can be extracted from N images captured by the same device. Denoting $\mathrm{W}^{(1)}, \mathrm{W}^{(2)}, \ldots, \mathrm{W}^{(N)}$ as the noise residue which is obtained by $\mathrm{I}^{(1)}, \mathrm{I}^{(2)}, \ldots, \mathrm{I}^{(N)}$ going through filter F, $\mathrm{W}^{(i)} = \mathrm{I}^{(i)} - F(\mathrm{I}^{(i)}), i = 1, \ldots, N$. the maximum likelihood estimator of PRNU noise $\hat{\mathbf{K}}$ is deduced according to the following formula [17], [18]:

$$\hat{\mathbf{K}} = \frac{\sum_{i=1}^{N} \mathrm{W}^{(i)}\mathrm{I}^{(i)}}{\sum_{i=1}^{N} \left(\mathrm{I}^{(i)}\right)^2}. \tag{2}$$

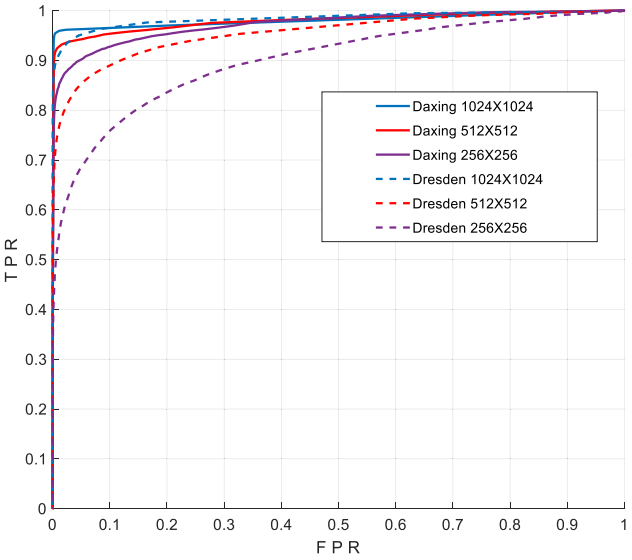


FIGURE 2. ROC performance of the image source identification algorithm [18] on Daxing dataset.

The fingerprint is estimated in the same way by the available images or video frames. Then, the Peak to Correlation Energy (PCE) [18] between the noise residue and the device fingerprint $\hat{\mathbf{K}}$ of is computed and compared to a threshold. If the PCE is higher than the threshold, then it is decided that fingerprint and noise residue have the same origin.

B. IMAGE SOURCE IDENTIFICATION

Image source identification, where a query image is matched with a device reference computed from a set of images taken by the device. In this scenario, the reference PRNU for each

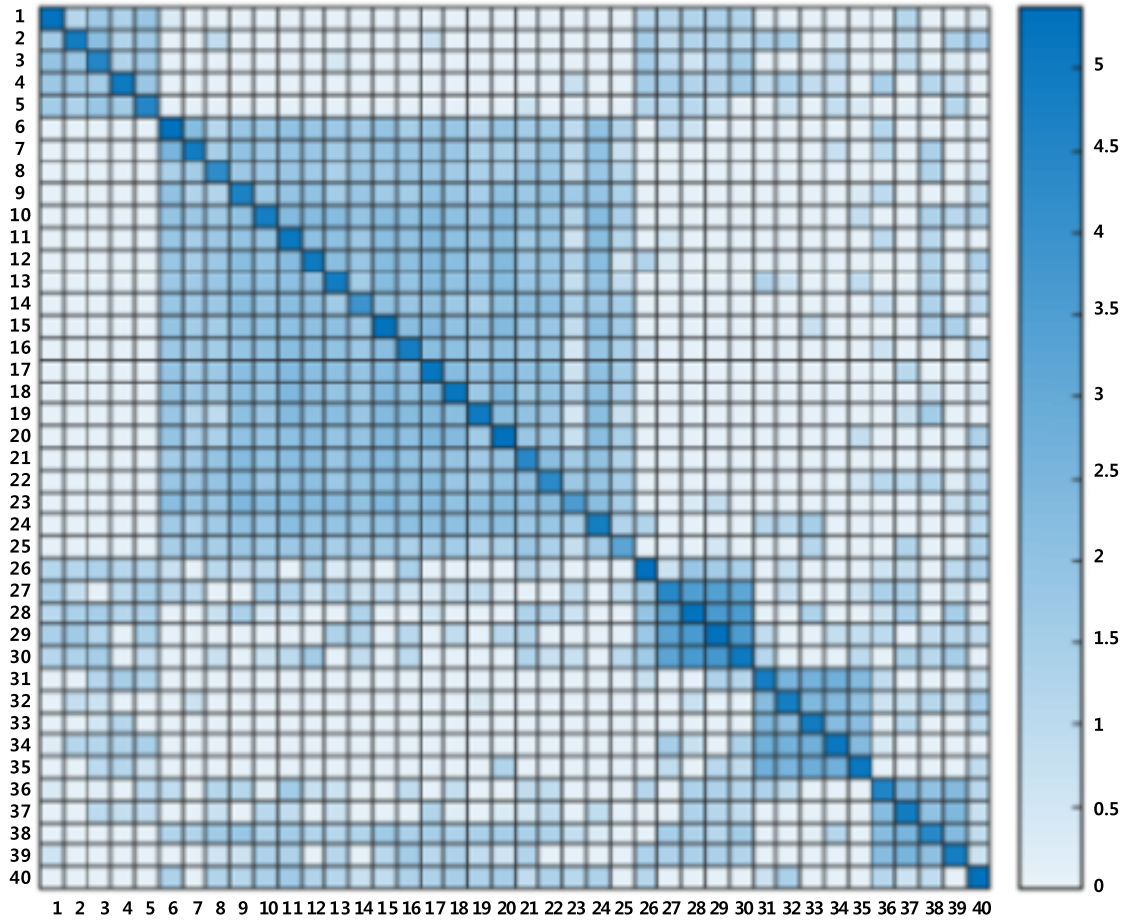


FIGURE 3. max(log(PCE),0) between 40 smartphones in Daxing Dataset. Huawei P10 Plus: 1-5, iPhone 6: 6-10, iPhone 6s: 11-15, iPhone 6s Plus: 16:20, iPhone 7 Plus: 21-25, OPPO R11: 26-:30, VIVO X9 (Plus): 31-35, Xiaomi 4A: 36-40.

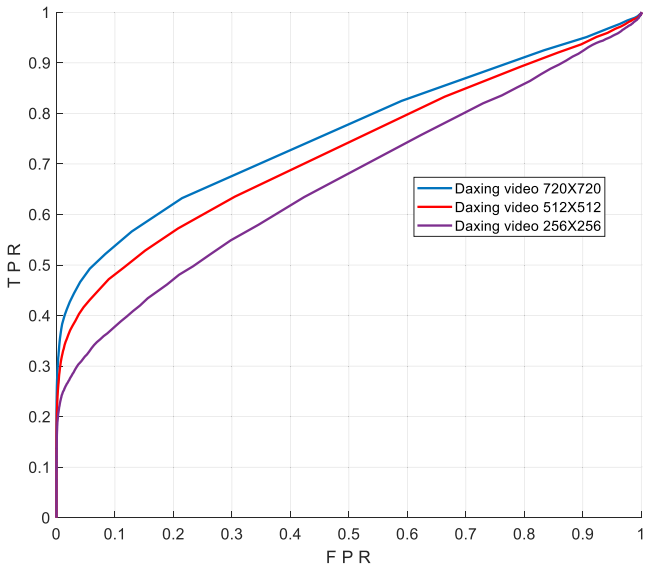


FIGURE 4. ROC performance of the video source identification algorithm [18] on Daxing dataset.

device is estimated from 50 images. Then, we carry out 3 experiments using 1024×1024, 512×512, and 256×256 images as queries, respectively. In all experiments, we set each device for 100 matching cases (images from the same smartphone) and the same number of mismatching cases (images randomly chosen from other smartphones). The experimental results are reported using ROC curves that plot true positive rate against false positive rate as shown in Fig. 2. In contrast, we also carry out the same experiments on Dresden dataset.

To evaluate the distinguishability between smartphones of the same model, we address 8 models (including Huawei P10 Plus, iPhone 6, iPhone 6S, iPhone 6S Plus, iPhone 7 Plus, OPPO R11, VIVO X9, and Xiaomi 4A) from the Daxing dataset. 5 devices of them are selected for each model, with a total of 40 smartphones. The $max\,(log\,(\mathrm{PCE})\,,0)$ values between these 40 smartphones are shown in Fig.3. As shown in Fig.3, the $max\,(log\,(\mathrm{PCE})\,,0)$ values between the same model of smartphones are higher than those between different models. It also implies that PRNU noise $\hat{\mathbf{K}}$ and noise residue $\mathrm{W}^{(i)}$, which are used for calculating PCE values, contain other noises related to camera models.

C. VIDEO SOURCE IDENTIFICATION

Video source identification, where a query frame is matched with a device reference computed from a set of frames

TABLE 3. Code and image resolution of each smartphone and its corresponding folders. Take code "2303" as an example, "2" refers to the Apple brand, "3" refers to the smartphone model of the iPhone 6s, "03" refers to the third iPhone 6S device.

Smartphone	Code	Resolution of Image	Resolution of Video	Processer	Operating System
Huawei P20 (1-5)	1101-1105	2976×3968	1920×1080	Kirin 970	ANDROID 9.0
Huawei Mate 9 (1-3)	1201-1203	2976×3968	1920×1080	Kirin 960	ANDROID 7.0
Huawei Mate 9 Plus (1)	1301	2976×3968	1920×1080	Kirin 960	ANDROID 7.0
Huawei P9 (1-5)	1401-1405	2976×3968	1920×1080	Kirin 955	ANDROID 7.0
Huawei P10 (1-3)	1501-1503	2976×3968	1920×1080	Kirin 960	ANDROID 7.0
Huawei 10 Plus (1-6)	1601-1606	2976×3968	1920×1080	Kirin 960	ANDROID (7.0, 7.0, 8.0, 8.0, 8.0, 8.0)
iPhone 6 (1-5)	2101-2105	2448×3264	1920×1080	iPhone A8	IOS (8, 11, 11, 11, 8)
iPhone 6 Plus (1-4)	2201-2204	3024×4032	1920×1080	iPhone A8	IOS (11, 10, 11, 11)
iPhone 6S (1-13)	2301-2313	3024×4032	1920×1080	iPhone A9	IOS (11, 11, 11, 11, 11, 11, 9, 11, 9, 10, 10, 10, 10)
iPhone 6S Plus (1-10)	2401-2410	3024×4032	1920×1080	iPhone A9	IOS (9, 9, 11, 11, 11, 11, 11, 11, 11, 11)
iPhone 7 (1-3)	2501-2503	3024×4032	1920×1080	iPhone A10	IOS 11
iPhone 7 Plus (1-5)	2601-2605	3024×4032	1920×1080	iPhone A10	IOS 11
iPhone 8 Plus (1-2)	2701-2702	3024×4032	1920×1080	iPhone A11	IOS (11, 12)
OPPO R9 (1-6)	3101-3106	3120×4160	1920×1080	MT 6755	ANDROID (5.1, 7.1, 6.0, 5.1, 5.1, 5.1)
OPPO R9S Plus (1)	3201	2592×4608	1920×1080	MSM 8976 Pro	ANDROID 7.1
OPPO R11 (1-5)	3301-3305	3456×4608	1920×1080	MSM 8976 Plus	ANDROID 7.1
OPPO R11T (1)	3401	3456×4608	1920×1080	MSM 8976 Plus	ANDROID 7.1
VIVO X9 (1-4)	4101-4104	3456×4608	1920×1080	MSM 8953	ANDROID (6.0, 7.1, 7.1, 7.1)
VIVO X9 Plus (1)	4201	3456×4608	1920×1080	MSM 8953	ANDROID 7.1
VIVO X9I (1)	4301	3456×4608	1280×720	MSM 8953	ANDROID 7.1
VIVO Y85 (1)	4401	3120×4160	1280×720	SDM 450	ANDROID 8.1
Xiaomi 4A (1-5)	5101-5105	3120×4160	1280×720	MSM 8917	ANDROID 6.0

of videos taken by the same device. Here, the reference PRNU for each smartphone is determined based on the first 50 frames of a video. Then, we carry out 3 experiments using 720×720, 512×512, and 256×256 images cropped from original frames as queries, respectively. In all experiments, we consider for each smartphone 100 matching cases (videos from the same smartphone) and the same number of mismatching cases (videos randomly chosen from other smartphones). The achieved results are reported using ROC curves that plot true positive rate against false positive rate as shown in Fig. 4.

V. CONCLUSION

To explore the personally identifiable information, in this paper, we build a Daxing smartphone identification dataset to focus on source forensics techniques for individual camera device identification. The dataset includes multiple smartphones from the same brand and model, and 43,400 original images and 1,400 original videos, from 90 smartphones of 22 models belonging to 5 brands. To the best of our knowledge, the Daxing dataset uses the largest amount of smartphones for image/video source identification compared with other related datasets, as well as the highest numbers of devices per model and captured images/videos. It is further beneficial to the criminal investigation and the critical forensic evidence. In the future, we will continue to collect more images and videos of smartphones to expand the dataset. In addition, we plan to add audio samples from these phones for smartphone identification.

ACKNOWLEDGMENT

F. A. Author thanks Yunfei Hao, Xinze Hao, Duo Yang, Yuxin Mao*et al.* for their hard work during the process of data collection and collation, and also thanks Rongrng Ni, and Pengpeng Yang for beneficial discussion and suggestion.

REFERENCES

[1] G. Schaefer and M. Stich, "UCID: An uncompressed color image database," *Proc. SPIE*, vol. 5307, pp. 472–480, Dec. 2003.

[2] T. Gloe and R. Böhme, "The dresden image database for benchmarking digital image forensics," *J. Digit. Forensic Pract.*, vol. 3, nos. 2–4, pp. 150–159, Feb. 2010.

[3] T. Gloe and R. Böhme, "The 'Dresden Image Database for benchmarking digital image forensics," in *Proc. 25th Symp. Appl. Comput.*, New York, NY, USA, Mar. 2010, pp. 1584–1590.

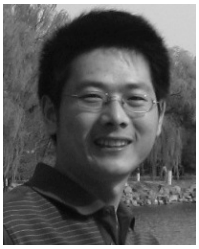
[4] D.-T. Dang-Nguyen, C. Pasquini, V. Conotter, and G. Boato, "RAISE: A raw images dataset for digital image forensics," in *Proc. 6th ACM Multimedia Syst. Conf.*, New York, NY, USA, Mar. 2015, pp. 219–224.

[5] D. Shullani, M. Fontani, M. Iuliani, O. A. Shaya, and A. Piva, "VISION: A video and image dataset for source identification," *EURASIP J. Inf. Secur.*, vol. 2017, no. 1, p. 15, Oct. 2017.

[6] *IEEE's Signal Processing Society—Camera Model Identification*. Accessed: Jul. 24, 2019. [Online]. Available: https://www.kaggle.com/c/sp-society-camera-model-identification

[7] O. A. Shaya, P. Yang, R. Ni, Y. Zhao, and A. Piva, "A new dataset for source identification of high dynamic range images," *Sensors*, vol. 18, no. 11, p. 3801, Nov. 2018.

[8] A. Abdelhamed, S. Lin, and M. S. Brown, "A high-quality denoising dataset for smartphone cameras," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Salt Lake City, UT, USA, Jun. 2018, pp. 1692–1700.
[9] C. Galdi, F. Hartung, and J.-L. Dugelay, "SOCRatES: A database of realistic data for source camera recognition on smartphones," in *Proc. Int. Conf. Pattern Recognit. Appl. Methods*, Feb. 2019, pp. 19–21.
[10] G. Qadir, S. Yahaya, A. T. S. Ho, "Surrey University library for forensic analysis (SULFA)," in *Proc. IET IPR*, London., U.K., 2012. pp. 1–6.
[11] Y. Hu, A. Salman, Y. Wang, B. Liu, and M. Li, "Construction and evaluation of video forgery detection database," *J. South China Univ. Technol., Natural Sci. Ed.,*, vol. 45, no. 12, pp. 57–64, 2017.
[12] J.-C. Li, Y.-J. Hu, A.-A. Mohammed, Y.-C. Xiong, D.-X. Wen, Y.-Y. Ren, and G.-J. Liao, "Expansion of video forgery detection database and validation of its effectiveness," *J. Appl. Sci.*, vol. 36, no. 2, pp. 347–361, Mar. 2018.
[13] L. Zheng, Y. Zhang, and L. Vrizlynn, "A survey on image tampering and its detection in real-world photos," *J. Vis. Commun. Image Represent.*, vol. 58, pp. 380–399, Jan. 2019.
[14] M. Zampoglou, S. Papadopoulos, and Y. Kompatsiaris, "Large-scale evaluation of splicing localization algorithms for Web images," *Multimedia Tools Appl.*, vol. 76, no. 4, pp. 4801–4834, Feb. 2017.
[15] G. Cao, Y. Zhao, R. Ni, and X. Li, "Contrast enhancement-based forensics in digital images," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 3, pp. 515–525, Mar. 2014.
[16] G. E. Healey and R. Kondepudy, "Radiometric CCD camera calibration and noise estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, no. 3, pp. 267–276, Mar. 1994.
[17] M. Chen, J. Fridrich, M. Goljan, and J. Lukas, "Determining image origin and integrity using sensor noise," *IEEE Trans. Inf. Forensics Security*, vol. 1, no. 1, pp. 74–90, Mar. 2008.
[18] M. Goljan, J. Fridrich, and T. Filler, "Large scale test of sensor fingerprint camera identification," *Proc. SPIE*, vol. 7254, Feb. 2009, Art. no. 72540I.



HUAWEI TIAN received the B.E. degree from the Chongqing University, in 2006, and the Ph.D. degree from the Institute of Information Science, Beijing Jiaotong University, in 2013. He is currently an Associate Professor with the People's Public Security University of China. His research interests include image processing, digital watermarking and steganography, digital forensics, and 3-D videos.

YANHUI XIAO received the B.S. and Ph.D. degrees from Beijing Jiaotong University, in 2007 and 2013, respectively. He is currently an Associate Professor with the People's Public Security University of China. His current research interests include digital forensics, intelligence analysis, face recognition, computer vision, and machine learning.

GANG CAO received the B.S. degree from the Wuhan University of Technology, Hubei, China, in 2005, and the Ph.D. degree from the Institute of Information Science, Beijing Jiaotong University, Beijing, China, in 2013. He was with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, in 2010. He is currently an Associate Professor with the School of Computer Science and Cybersecurity, Communication University of China, Beijing. His current research interests include digital forensics, data hiding, and multimedia signal processing.

YONGSHENG ZHANG received the B.E. degree from the Shandong Police College, in 2017. He is currently pursuing the M.S. degree with the School of Criminal Investigation and Counter Terrorism, People's Public Security University of China. His current research interests include image processing and digital forensics.

ZHIYIN XU is currently pursuing the B.S. degree with the School of Criminal Investigation and Counter Terrorism, People's Public Security University of China. His current research interests include digital forensics and watermarking.

YAO ZHAO received the B.S. degree from the Radio Engineering Department, Fuzhou University, Fuzhou, China, in 1989, and the M.E. degree from the Radio Engineering Department, Southeast University, Nanjing, China, in 1992, and the Ph.D. degree from the Institute of Information Science, Beijing Jiaotong University (BJTU), Beijing, China, in 1996.

He became an Associate Professor at BJTU, in 1998, and became a Professor, in 2001. From 2001 to 2002, he was a Senior Research Fellow with the Information and Communication Theory Group, Faculty of Information Technology and Systems, Delft University of Technology, Delft, The Netherlands. In 2015, he visited the Swiss Federal Institute of Technology, Lausanne, Switzerland. From 2017 to 2018, he visited the University of Southern California. He is currently the Director of the Institute of Information Science, BJTU. His current research interests include image/video coding, digital watermarking and forensics, and video analysis and understanding.

Dr. Zhao serves on the Editorial Boards of several international journals, including as an Associate Editor of the IEEE TRANSACTIONS ON CYBERNETICS, and the IEEE SIGNAL PROCESSING LETTERS, and an Area Editor of *Signal Processing: Image Communication*. He was named a Distinguished Young Scholar by the National Science Foundation of China, in 2010, and elected as a Chang Jiang Scholar of Ministry of Education of China, in 2013. He is a Fellow of the IET.