SOCRatES: A Database of Realistic Data for SOurce Camera REcognition on Smartphones

Chiara Galdi[1], Frank Hartung[2] and Jean-Luc Dugelay[1]
[1]*Department of Digital Security, EURECOM, 450 Route des Chappes, 06410 Biot, France*
[2]*Department of Multimedia Technology, FH Aachen, Eupener Str. 70, 52066 Aachen, Germany*

Keywords: Sensor Pattern Noise, PRNU, Source Camera Identification, Video, Smartphone, Database.

Abstract: SOCRatES: SOurce Camera REcognition on Smartphones, is an image and video database especially designed for source digital camera recognition on smartphones. It answers to two specific needs, the need of wider pools of data for the developing and benchmarking of image forensic techniques, and the need to move the application of these techniques on smartphones, since, nowadays, they are the most employed devices for image capturing and video recording. What makes SOCRatES different from all previous published databases is that it is collected by the smartphone owners themselves, introducing a great heterogeneity and realness in the data. SOCRatES is currently made up of about 9.700 images and 1000 videos captured with 103 different smartphones of 15 different makes and about 60 different models. With 103 different devices, SOCRatES is the database for source digital camera identification that includes the highest number of different sensors. In this paper we describe SOCRatES and we present a baseline assessment based on the Sensor Pattern Noise computation.

1 INTRODUCTION

It is a fact that, nowadays, the most employed devices for recording videos and capturing photos are smartphones. The high resolution provided by their embedded acquisition sensors allows the recording of amateur videos or pictures of good quality. This has spread their use for collecting souvenir photos, replacing the classic cameras, but also in collecting covert videos and illegal contents, including pedopornography, bullying, and illegal races. In the latter case, it is extremely important to have tools to reliably associate an image or a video with illegal content to the correct source camera. This research field is referred to as source digital camera identification.

For the aforementioned reasons, the authors present in this paper a novel database, namely SOCRatES: SOurce Camera REcognition on Smartphones. This image and video database is especially designed for the purpose of development and benchmarking of image forensic techniques on smartphones, in particular for, but not limited to, the source digital camera identification problem.

1.1 Source Digital Camera Identification

The problem of source digital camera identification has been addressed in different ways during the last decades. Three main categories of approaches can be distinguished:

The first category consists in approaches based on analysing the artefacts produced in the acquisition phase. The **lens aberration** is an optical property that causes light passing through a lens to be spread out over some region of space rather than focused to a point. The consequent radial distortion causes straight lines to appear as curved on the output image. The image distortion is analysed to identify the source camera. This approach was first proposed by Choi et al. in (Choi et al., 2006). Imperfections in the lens may also produce **chromatic aberration**. The latter has been studied by Van and al. in (Van et al., 2007). However, cameras of the same model or mounting the same lens system will produce the same distinctive pattern. This method is thus suitable for camera model identification but not for distinguishing cameras of the same model.

The second category includes approaches able to uniquely link the captured image to its source cam-

Galdi, C., Hartung, F. and Dugelay, J.
SOCRatES: A Database of Realistic Data for SOurce Camera REcognition on Smartphones.
DOI: 10.5220/0007403706480655
In *Proceedings of the 8th International Conference on Pattern Recognition Applications and Methods (ICPRAM 2019)*, pages 648-655
ISBN: 978-989-758-351-3

era as it analyses the sensor imperfections, also called **sensor noise**, that are different for each camera even if they are of the same make and model. The sensor noise is the result of three main components: pixel defects, fixed pattern noise (FPN), and Photo Response Non Uniformity (PRNU). Two methods fall in this category. The analysis of **pixel defects** consists in examine the defects of Charge-Coupled Device (CCD) sensors, including point defects, hot point defects, dead pixel, pixel traps, and cluster defects. To extract the pixel defect pattern, pictures of a black surface must be taken. Some limitations of this method include the sensor sensitivity to temperature that may affect the extracted pattern, the image content can make the pattern less visible, the pattern changes with time as the sensor ages (e.g. the number of defective pixels increases). Finally, for some cameras it is possible that they do not have any defective pixels and thus a distinctive pattern. Such approach is presented in (Geradts et al., 2001) by Geradts et al.

The second method in the sensor imperfections category is also the one dealt with in this article. The **sensor pattern noise** (SPN) is a distinctive pattern due to imperfections in the silicon wafer during the sensor manufacturing, different even among cameras of the same model. These imperfections imply that the pixels have different sensitivities to light. A distinctive pattern can be extracted by analysing the image in the frequency domain and by selecting those frequencies that are more likely to be associated with the sensor noise. The method was first presented by Lukas et al. in (Lukas et al., 2006) and is described in more details in Section 3. This method is widely adopted and research is still very active in this field, as there are some open issues: the model assumes that the reference pattern and the test image have the same size, thus the method fails to predict the source camera of cropped images; strong image or video compression, such as that applied to media files when uploaded to social networks, impact the noise pattern and produce a loss of accuracy.

The third category includes methods based on the analysis of the traces left on the image by the processing performed by the imaging device. The camera colour filter array (CFA) is a mosaic of tiny colour filters placed over the pixel sensors of an image sensor to capture colour information. The **CFA interpolation** process leaves a trace on the image and different approaches have been developed to extract a distinctive pattern from it. The methods include: to examine the traces of colour interpolation in colour bands, quadratic pixel correlation model, and binary similarity measures. As different cameras can adopt the same CFA interpolation, these approaches are suitable for make or model recognition rather than to uniquely associate the image to its source camera.

The works presented in (Lanh et al., 2007) and (Redi et al., 2011) provide an extensive survey on digital camera image forensics.

1.2 Databases for Source Digital Camera Identification

A broad range of scientific literature exists in the field of the digital image forensic. However, there are very few publicly available database especially designed for source camera identification.

The first large and publicly available image database has been proposed in 2010, namely the 'Dresden Image Database' (Gloe and Böhme, 2010). This database is composed by more than 14,000 images acquired with 73 different cameras of 25 different models and is intended for developing and benchmarking of camera model identification techniques. It has been used in a number of works, including the recent work on camera model identification based on local features by Marra et al. (Marra et al., 2017), the work presented by Deng et al. in (Deng et al., 2011) where the authors propose a new technique based on the approximation of the Auto-White Balance algorithm used inside cameras. It has also been used in combination with a custom-made dataset in order to have a wider benchmark, as in (Lin and Li, 2016). The Dresden Image Database has been also employed by Gloe et al. (Gloe et al., 2012) to analyse unexpected artefacts in PRNU-based digital camera identification.

Another small database for blind source cellphone model identification has been presented in 2008 by Çeliktutan et al. in (Çeliktutan et al., 2008). It contains more than 3.000 pictures collected using 17 mobile phones of 15 different models. In a work proposed by Farinella et al. (Farinella et al., 2015) and published in 2015, this database is used in combination with the 'Dresden Image Database' to compare two known techniques for source camera identification, namely the one based on sensor pattern noise extraction and the approach based on the analysis of the specific colour processing and transformations operated by the camera before storing. The fact that a work presented in 2015 had to employ two databases collected in 2008 and in 2010, brings into focus the necessity of having more and up-to-date image databases. This is the case in particular for image databases collected with mobile phones, since smartphone features rapidly improve over time, for example, in the database collected in 2008, the maximum capturing resolution is of 640×480 pixels while in SOCRatES the maximum resolution is of 5344×3006 pixels.

More recently, the VISION database has been release and presented in (Shullani et al., 2017). The database is composed by 34,427 images and 1914 videos, both in the native format and in their social version (Facebook, YouTube, and WhatsApp), from 35 portable devices of 15 major brands. It has been recently used to test CNN-based techniques and to investigate their vulnerability to adversarial attacks for camera model identification (Marra et al., 2018).

The advantages offered by SOCRatES are twofolded, it currently offers the highest number of different sensors employed for data collection, and it is the only database presented so far for digital source camera identification collected by the smartphone owners themselves, introducing a great heterogeneity and realness in the data.

SOCRatES is particularly designed for testing approaches based on the Sensor Pattern Noise extraction, e.g. the technique firstly presented by Lukas et al. in (Lukas et al., 2006). It is currently made up of about 9.700 images and 1000 videos captured with 103 different smartphones of 15 different makes and about 60 different models.

The use of digital image forensic techniques is not limited to the investigation of crime, it has been also applied for user authentication by combining the smartphone identification with the user biometric recognition, in order to provide an easy-to-use and reliable authentication system. In (Galdi et al., 2016), smartphone identification is combined with iris recognition. The same authors have presented more recently, a method combining smartphone identification and face recognition using the SOCRatES database (Galdi et al., 2018).

In addition to photos, SOCRatES includes a set of video clips collected with each different device. The problem of source digital camera identification from strongly compressed videos, as the ones generated by smartphones, is very tough (Chuang et al., 2011) since the sensor pattern noise is strongly impacted by video compression. Also, compared to photos taken with the same device, the recorded scene is somehow cropped. This is why videos are included in SOCRatES, in order to provide a benchmark for testing techniques for source digital camera identification from videos on smartphones.

SOCRatES is made freely available to other researchers for scientific purposes at the following URL: `http://socrates.eurecom.fr/`.

The reminder of this paper is organized as follows: in section 2 the SOCRatES database is described, including its acquisition protocol, structure and annotation. Section 3 presents a preliminary analysis of the database, including source camera identification performances. Section 4 concludes the paper.

Figure 1: Guidelines for collecting uniform colour background pictures.

Figure 2: Guidelines for collecting pictures avoiding copyright and privacy violations. Some pictures have been obscured here for copyright reasons.

2 DATABASE DESCRIPTION

SOCRatES is currently made up of about 9.700 images and 1000 videos captured with 103 different smartphones of 15 different makes and about 60 different models. The acquisition has been performed in uncontrolled conditions. In order to collect the database, many people were involved and asked to use their personal smartphone to collect a set of pictures. Instructions were given to the participants and they collected the set of pictures in complete autonomy.

The reason behind this choice is, on the one hand, to collect a database of heterogeneous pictures and to maximize the number of devices employed, and, on the other hand, to carefully replicate the real scenario of application of the techniques that will use this database as benchmark. In fact, by selecting a "population" of smartphone users and letting them capturing the pictures, we automatically select a set of smartphones representing the current market.

Table 1 summarizes the main features of the de-

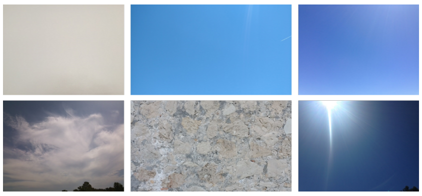

Figure 3: Sample images from the background pictures set.

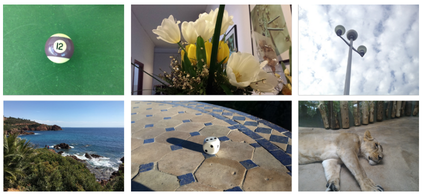

Figure 4: Sample images from the foreground pictures set.

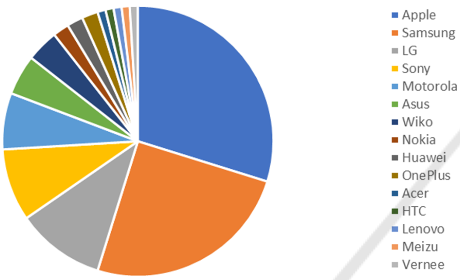

Figure 5: Smartphone brands composing SOCRatES.

vices composing the database.

2.1 Acquisition Protocol

Participants are asked to use their personal smartphone and to collect a total of 100 photos and 10 videos. Among the 100 captured pictures: 50 have to be photos of the blue sky, or, in its absence, of another uniform colour surface, e.g. a white wall; 50 pictures have to portray any kind of scene, avoiding privacy and copyright sensitive subjects, e.g. faces, people, copyrighted buildings, license plates, brands, etc. In Fig. 1 and Fig. 2 an example of the illustrated indications given to the participants as a guideline for capturing the photos is shown. All pictures are then checked by the database owners to ensure that they do not violate privacy and copyright. All volunteers are informed about the purpose of the data collection and assignments of copyright were signed by all those willing to help.

Participants are also asked to set the camera to the maximum available resolution, to capture photos in landscape format, i.e. horizontally, to always use the same camera, preferably the one with the best resolution that is usually the rear camera.

2.2 Background Pictures

With the term "background pictures" we indicate the subset of photos portraying uniform colour scene, preferably the blue sky. This kind of pictures were included in the database since they are used in a number of techniques in order to extract the sensors reference pattern noise, e.g. in the approaches presented in (Lukas et al., 2006) (Li, 2009).

However, since the pictures were captured by different persons, and despite the instructions given to the participants, some of the background pictures portray non-uniform backgrounds. Some sample background images are shown in Fig. 3, where the first row represents "good" background pictures and the second row the "bad" ones.

2.3 Foreground Pictures

We employ the term "foreground pictures" to indicate the photos portraying any kind of scene and in opposition to the "background pictures". These pictures are very heterogeneous, since they were captured by different devices, by different persons, in different places, and at different times. Some sample foreground images are shown in Fig. 4.

2.4 Videos

Ten short video clips are recorded with each device. Their duration varies from 2 to 10 seconds.

2.5 Database Structure and Annotation

A naming convention has been adopted to distinguish the images/videos captured with different devices, an ID number has been assigned to each different device, and to indicate the type of the acquired item, i.e.: "background picture", "foreground picture", "video".

Along with pictures and videos, annotation files describing the characteristics of the smartphones employed are provided. In particular they list, for each device, the smartphone model, the Operating System, the digital camera model, the photo resolution and the video resolution employed during acquisition.

In Fig. 5, a graph of the distribution of the 60 different smartphone brands included in SOCRatES is given.

The database is released under a license agreement ensuring the compliance with the current European regulations. Researchers shall use the Database only for non-commercial research and educational purposes.

Table 1: Devices main features. IR = Image Resolution; #BG = number of background pictures; #FG = number of foreground pictures; VR = Video Resolution.

ID	Brand	Model	IR	#BG	#FG	VR	#video
100	Motorola	X Play	5344x3006	50	40	1920x1080	10
101	Samsung	Galaxy S5 (SM-G900F)	5312x2988	50	40	1920x1080	10
102	LG	G3 D855	4160x2340	50	40	1920x1080	10
104	Samsung	Galaxy S5	5312x2988	50	40	1920x1080	10
105	Wiko	Birdy 4G	2560x1920	50	40	1920x1088	10
107	Apple	iPhone 6	3264x2448	50	40	1920x1080	10
108	Apple	iPhone 6	3264x2448	50	40	1920x1080	10
109	Apple	iPhone 6	3264x2448	50	40	1920x1080	10
110	Huawei	P8 Lite	4160x3120	50	40	1920x1088	10
111	LG	G3	4160x3120	50	40	1920x1080	10
112	Motorola	Moto G (XT1072)	3264x1836	50	50	1280x720	10
113	Sony	E6653	3840x2160	50	40	1920x1080	10
114	Apple	iPhone 6s	4032x3024	50	40	1920x1080	10
115	Samsung	Galaxy Core Prime	2592x1944	50	40	1280x720	10
116	LG	G4	5312x2988	50	40	-	0
117	Acer	Liquid E700	3840x2160	50	40	1920x1088	10
118	Nokia	Lumia 635	1920x1080	50	40	1280x720	10
119	Wiko	Rainbow 4G	3264x2448	50	40	1280x720	10
120	Apple	iPhone 5c	3264x2448	50	40	1920x1080	10
121	Motorola	Moto G	2592x1944	50	40	-	10
123	Samsung	Galaxy S6 Edge	5312x2988	50	46	640x368	10
124	Samsung	Galaxy S3 Neo (GT-i9301i)	1280x720	50	40	1920x1080	10
125	Huawei	P7	4160x2336	50	40	1280x720	10
126	LG	Nexus 5	3264x2448	50	40	1920x1080	10
127	Sony	Xperia Z1 Compact	3840x2160	50	40	1920x1080	10
128	Apple	iPhone 6s	4032x3024	50	40	1920x1080	10
129	Apple	iPhone 5c	3264x2448	50	50	1920x1080	10
130	Lenovo	S60-a	4096x2304	50	40	1920x1080	10
131	Samsung	Galaxy S3 Neo (GT-i9301i)	3264x1836	50	40	1920x1080	11
132	Motorola	Moto X-Style	5344x3006	50	40	1920x1080	10
133	Samsung	Note 4	5312x2988	50	40	1920x1080	10
135	Samsung	Galaxy Grand Prime	3264x2448	50	40	1920x1080	10
136	Apple	iPhone 4s	3264x1836	50	50	1920x1080	10
137	Apple	iPhone 6	3264x2448	50	50	1920x1080	10
138	Sony	Xperia Z3	1278x718	50	50	1920x1080	10
139	Samsung	Galaxy Core Max (SM-G5108Q)	3264x1836	50	40	1920x1080	10
140	LG	G3 D855	4160x2340	50	40	1920x1080	10
141	Asus	Zenfone 2	4096x3072	50	40	1920x1080	10
142	Apple	iPhone 5c	3264x2448	50	40	1920x1080	10
143	Sony	Xperia Z3	3840x2160	50	40	1920x1080	10
144	HTC	One M8	2688x1520	50	40	1920x1080	10
145	Asus	Zenfone 2 (ZE551ML)	4096x3072	50	40	1920x1080	5
146	Apple	iPhone 4s	3264x2448	50	40	1920x1080	10
147	Motorola	Moto G	4160x2340	50	40	1920x1080	10
148	Apple	iPhone 5s	3264x2448	50	40	1920x1080	10
149	LG	Spirit LTE	3264x1840	50	40	1280x720	10
150	Apple	iPhone 6	640x480	50	40	1280x720	10
152	Samsung	Galaxy Grand Plus	2560x1536	50	40	1280x720	10
154	Apple	iPad mini 2	1026x766	50	39	1920x1080	10
155	OnePlus	X	4160x3120	50	40	1920x1080	10
156	Sony	M4	3920x2204	50	40	1920x1080	10

Table 1. (continuation)

ID	Brand	Model	IR	#BG	#FG	VR	#video
157	Sony	Xperia Z (C6603)	3920x2204	50	40	1920x1080	10
158	Samsung	Galaxy Core Prime	2560x1536	50	40	1280x720	10
159	Samsung	Galaxy S6	5312x2988	50	40	1920x1080	10
160	OnePlus	One	4160x3120	50	40	3840x2160	10
161	Samsung	Galaxy S5 Mini	3264x1836	50	40	1920x1080	10
162	Samsung	Galaxy S4	4128x2322	50	40	1920x1080	10
163	Samsung	A510F	4128x2322	50	50	1920x1080	10
165	Apple	iPhone 5s	3264x2448	50	40	1920x1080	10
166	Samsung	Galaxy S5	5312x2988	50	40	-	13
167	Apple	iPhone 5c	2049x1536	50	40	1280x720	10
168	Wiko	Highway 4G	640x480	50	40	1280x720	10
169	Apple	iPhone 6	3264x2448	50	50	1920x1080	10
170	Apple	iPhone 5c	3264x2448	50	50	1920x1080	10
171	Asus	Zenfone MAX	3024x4032	50	50	1920x1080	10
172	Apple	iPhone 7	4032x3024	50	50	1920x1080	10
173	Apple	iPhone SE	4032x3024	50	50	1920x1080	10
174	Samsung	Galaxy S5	4608x2592	50	50	1920x1080	10
175	Apple	iPhone 6s plus	4032x3024	50	50	1920x1080	10
176	Sony	Xperia Z3 (D6603)	3840x2160	50	50	1920x1080	10
177	Apple	iPhone 4s	3264x2448	50	50	1920x1080	10
179	Apple	iPhone 5c	3264x2448	50	50	1920x1080	10
182	Samsung	Galaxy S3	4128x3096	52	53	1920x1080	10
183	Apple	iPhone 7	4032x3024	50	50	3840x2160	10
185	Samsung	Galaxy S7 Edge	4032x3024	50	50	1920x1080	10
186	LG	Nexus 5X	4032x3024	50	51	1920x1080	10
187	LG	Nexus 5X	4032x3024	50	50	3840x2160	10
189	Samsung	Galaxy A3 (2016)	4128x2322	50	50	1920x1080	10
190	Apple	iPhone 7	4032x3024	50	50	1920x1080	10
191	Asus	Zenfone 3	4096x2304	50	50	1280x720	10
193	Vernee	Thor	4864x2736	50	50	1280x720	10
194	Sony	Xperia T3	3104x1746	50	50	1920x1080	20
195	Apple	iPhone 6	3264x2448	50	50	1280x720	10
196	Samsung	Galaxy A3 (2016)	3264x2448	50	50	1920x1080	10
197	Meizu	M3 Note	2560x1440	50	50	1920x1080	10
198	Motorola	Moto G3	4160x2340	50	50	1920x1080	10
199	LG	G4	5312x2988	50	50	1920x1080	10
200	Wiko	Rainbow Up 4G	3264x2448	48	50	-	0
201	Apple	iPhone 6s	4032x3024	50	50	1920x1080	11
202	Samsung	Galaxy S4	4128x2322	60	40	1920x1080	10
204	Nokia	Lumia 930	3552x2000	50	40	1920x1080	10
210	Apple	iPhone 6	3264x2448	50	50	1920x1080	10
211	Apple	iPhone 5	960x720	50	50	576x320	10
212	Asus	Zenfone 2 (ZE551ML)	4096x2304	50	53	1920x1080	13
213	Sony	Xperia E3	2560x1440	50	50	1920x1080	10
214	Samsung	Galaxy J7	4128x2322	50	50	1920x1080	10
215	Apple	iPhone 6	3264x2448	50	50	1920x1080	10
216	LG	K10 4G	4160x2340	50	50	1280x720	10
217	Motorola	Moto G3	4160x2340	50	50	1920x1080	10
219	Samsung	Galaxy J7 2016	4128x3096	50	50	1920x1080	10
220	Samsung	Galaxy S4 mini	1280x720	50	50	1920x1080	10
224	LG	G3	2048x1536	50	50	3840x2160	10
225	Samsung	S7 Edge	4032x3024	73	50	1920x1080	10

3 BASELINE ASSESSMENT

In this section, the baseline assessment based on two well-known techniques, presented in (Lukas et al., 2006) and (Li, 2009), is reported. The purpose of this evaluation is to provide the researchers willing to use this database with a starting point to be used for comparisons in the evaluation of their techniques. The analysis is based on the extraction and comparison of the Sensor Pattern Noise (SPN in the following) (Lukas et al., 2006). The SPN can be seen as the sensor "fingerprint", a distinctive pattern due to imperfections in the silicon wafer during the sensor manufacturing, different even among cameras of the same model. The SPN n is computed as follows:

$$n = DWT(I) - F(DWT(I))$$

where $DWT()$ is the discrete wavelet transform to be applied on image I and $F()$ is a denoising function applied in the DWT domain. For a more detailed description of $F()$, the reader is referred to appendix A of (Lukas et al., 2006).

For each device, its Reference SPN (RSPN) is computed using its "background pictures". The RSPN n_r corresponds to the average SPN computed over N images:

$$n_r = \frac{1}{N} \times \sum_{k=1}^{N} n_k$$

In order to test if a picture comes from a given device, its SPN is compared with the device RSPN. The higher the correlation, the more likely the photo comes from the device. Correlation is computed as follows:

$$corr(n, n_r) = \frac{(n - \bar{n}) * (n_r - \bar{n_r})}{||n - \bar{n}||||n_r - \bar{n_r}||}$$

where the bar above a symbol denotes the mean value.

3.1 Lukas et al.'s Approach Performance on SOCRatES

The performances of the method proposed by Lukas et al. in (Lukas et al., 2006), are summarized in this section. The RSPN is extracted, as described above, from the "background pictures" for each device using the code made publicly available by the authors[1]. Then the SPN is computed for each "foreground picture" and associated to the most correlated RSPN.

Performances are assessed in terms of Equal Error Rate (EER), Receiver Operating Characteristic curve (ROC) and Area Under the ROC curve, and summarised in Table 2.

[1] http://dde.binghamton.edu/download/camera_fingerprint/

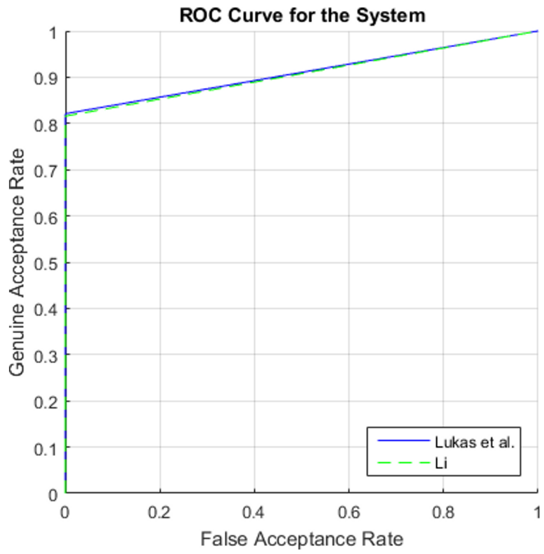

Figure 6: ROC curve illustrating the baselines assessment on SOCRatES.

Table 2: Performances of Lukas et al. and Li on SOCRatES.

	EER	AUC	RR
Lukas et al.	0.0894	0.9106	0.9191
Li	0.0921	0.9079	0.9164

3.2 Li's Approach Performance on SOCRatES

Li's approach proposes an enhancing process to mitigate the impact of scene details in the computation of the SPN. The Enhanced SPN (ESPN in the following) n_e is obtained as follows:

$$n_e(i,j) = \begin{cases} e^{-0.5n^2(i,j)/\alpha^2}, & \text{if } 0 <= n(i,j) \\ -e^{-0.5n^2(i,j)/\alpha^2}, & \text{otherwise} \end{cases}$$

where n_e is the ESPN, n is the SPN, i and j are the indices of the components of n and n_e, and α is a parameter that is set to 7, as indicated in (Li, 2009).

As in the first experiment, the RSPN is extracted from the "background pictures". Then the ESPN is computed for each "foreground picture" and associated to the most correlated RSPN, i.e. each "foreground picture" is associated to the most correlated camera.

The ROC curves obtained by the two tested methods are compared in Fig. 6.

4 CONCLUSION

One of the most important contributions to Western thought of the classical Greek philosopher Socrates is his dialectic method of inquiry, which is the foundation of the modern scientific method. This is why we found his name appropriate for a database designed for image and video forensic.

SOCRatES is a publicly available database intended for source digital camera identification on smartphones. In other fields, several databases are merged together to have a wider pool of data. This is done in particular for developing and benchmarking of deep-learning based techniques that require thousands of images and are the trend at the moment. SOCRatES can be used alone or in combination with other image or video databases in order to widen the data pool. Also, its challenging data samples, make it very suitable as testing set.

In this paper the SOCRatES database is described and baseline performances are obtained by testing two well-known techniques based on the Sensor Pattern Noise computation. The latter is a technique to identify, given a picture, its source digital camera. In particular, this technique can distinguish devices of the same make and model.

Another important feature of SOCRatES, is the presence of both images and videos captured with each device. This allows the study of source camera recognition on strongly compressed videos, which is still an open issue, as for the study of asymmetric comparison between videos and still images.

SOCRatES is made freely available to other researchers for scientific purposes at the following URL: http://socrates.eurecom.fr/.

REFERENCES

Çeliktutan, O., Sankur, B., and Avcibas, I. (2008). Blind identification of source cell-phone model. *IEEE Trans. Information Forensics and Security*, 3(3):553–566.

Choi, K. S., Lam, E. Y., and Wong, K. K. (2006). Source camera identification using footprints from lens aberration. In *Digital Photography II*, volume 6069, page 60690J. International Society for Optics and Photonics.

Chuang, W.-H., Su, H., and Wu, M. (2011). Exploring compression effects for improved source camera identification using strongly compressed video. In *Image Processing (ICIP), 2011 18th IEEE International Conference on*, pages 1953–1956. IEEE.

Deng, Z., Gijsenij, A., and Zhang, J. (2011). Source camera identification using auto-white balance approximation. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 57–64. IEEE.

Farinella, G. M., Giuffrida, M. V., Digiacomo, V., and Battiato, S. (2015). On blind source camera identification. In *International Conference on Advanced Concepts for Intelligent Vision Systems*, pages 464–473. Springer.

Galdi, C., Nappi, M., and Dugelay, J.-L. (2016). Multimodal authentication on smartphones: Combining iris and sensor recognition for a double check of user identity. *Pattern Recognition Letters*, 82:144–153.

Galdi, C., Nappi, M., Dugelay, J.-L., and Yu, Y. (2018). Exploring new authentication protocols for sensitive data protection on smartphones. *IEEE Communications Magazine*, 56(1):136–142.

Geradts, Z. J., Bijhold, J., Kieft, M., Kurosawa, K., Kuroki, K., and Saitoh, N. (2001). Methods for identification of images acquired with digital cameras. In *Enabling technologies for law enforcement and security*, volume 4232, pages 505–513. International Society for Optics and Photonics.

Gloe, T. and Böhme, R. (2010). The'dresden image database'for benchmarking digital image forensics. In *Proceedings of the 2010 ACM Symposium on Applied Computing*, pages 1584–1590. ACM.

Gloe, T., Pfennig, S., and Kirchner, M. (2012). Unexpected artefacts in prnu-based camera identification: a'dresden image database'case-study. In *Proceedings of the on Multimedia and security*, pages 109–114. ACM.

Lanh, T. V., Chong, K., Emmanuel, S., and Kankanhalli, M. S. (2007). A survey on digital camera image forensic methods. In *2007 IEEE International Conference on Multimedia and Expo*, pages 16–19.

Li, C.-T. (2009). Source camera identification using enahnced sensor pattern noise. In *Image Processing (ICIP), 2009 16th IEEE International Conference on*, pages 1509–1512. IEEE.

Lin, X. and Li, C.-T. (2016). Enhancing sensor pattern noise via filtering distortion removal. *IEEE Signal Processing Letters*, 23(3):381–385.

Lukas, J., Fridrich, J., and Goljan, M. (2006). Digital camera identification from sensor pattern noise. *IEEE Transactions on Information Forensics and Security*, 1(2):205–214.

Marra, F., Gragnaniello, D., and Verdoliva, L. (2018). On the vulnerability of deep learning to adversarial attacks for camera model identification. *Signal Processing: Image Communication*, 65:240–248.

Marra, F., Poggi, G., Sansone, C., and Verdoliva, L. (2017). A study of co-occurrence based local features for camera model identification. *Multimedia Tools and Applications*, 76(4):4765–4781.

Redi, J. A., Taktak, W., and Dugelay, J.-L. (2011). Digital image forensics: a booklet for beginners. *Multimedia Tools and Applications*, 51(1):133–162.

Shullani, D., Fontani, M., Iuliani, M., Al Shaya, O., and Piva, A. (2017). Vision: a video and image dataset for source identification. *EURASIP Journal on Information Security*, 2017(1):15.

Van, L. T., Emmanuel, S., and Kankanhalli, M. S. (2007). Identifying source cell phone using chromatic aberration. In *Multimedia and Expo, 2007 IEEE International Conference on*, pages 883–886. IEEE.