

Received January 12, 2022, accepted February 9, 2022, date of publication February 14, 2022, date of current version February 25, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3151406

A New Forensic Video Database for Source Smartphone Identification: Description and Analysis

YOUNES AKBARI¹, SOMAYA AL-MAADEED¹, (Senior Member, IEEE),
NOOR AL-MAADEED¹, (Member, IEEE), AL ANOOD NAJEEB¹,
AFNAN AL-ALI¹, FOUAD KHELIFI², (Member, IEEE),
AND ASHREF LAWGALY²

¹Department of Computer Science and Engineering, Qatar University, Doha, Qatar

²Department of Computer and Information Sciences, Northumbria University, Newcastle upon Tyne, NE1 8ST, U.K.

Corresponding author: Somaya Al-Maadeed (s_alali@qu.edu.qa)

This work was supported in part by the National Priorities Research Program (NPRP) through the Qatar National Research Fund (a member of Qatar Foundation) under Grant NPRP12S-0312-190332, and in part by the Open Access Funding Provided by the Qatar National Library.

ABSTRACT In recent years, the field of digital imaging has made significant progress, so that today every smartphone has a built-in video camera that allows you to record high-quality video for free and without restrictions. On the other hand, rapidly growing internet technology has contributed significantly to the widespread use of digital video via web-based multimedia systems and mobile smartphone applications such as YouTube, Facebook, Twitter, WhatsApp, etc. However, as the recording and distribution of digital videos have become affordable nowadays, security issues have become threatening and spread worldwide. One of the security issues is identifying source cameras on videos. There are some new challenges that should be addressed in this area. One of the new challenges is individual source camera identification (ISCI), which focuses on identifying each device regardless of its model. The first step towards solving the problems is a popular video database recorded by modern smartphone devices, which can also be used for deep learning methods that are growing rapidly in the field of source camera identification. In this paper, a smartphone video database named Qatar University Forensic Video Database (QUFVD) is introduced. The QUFVD includes 6000 videos from 20 modern smartphone representing five brands, each brand has two models, and each model has two identical smartphone devices. This database is suitable for evaluating different techniques such as deep learning methods for video source smartphone identification and verification. To evaluate the QUFVD, a series of experiments to identify source cameras using a deep learning technique are conducted. The results show that improvements are essential for the ISCI scenario on video.

INDEX TERMS Database, smart phone, source camera identification on videos, deep learning methods.

I. INTRODUCTION

Cellphone has developed rapidly over the past century due to its economic advantages, functionality and ease of access [1]. It allows the creation of digital audiovisual content without any constraints such as time, objects, places and network connections [2]. Smartphone devices can provide some pertinent information for crime prosecution and forensic investigations in massively important manner [1]. These types

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

of investigations have a potential importance in research fields in all the sectors like medicine, law, and surveillance system where images and videos authenticity is important. In general, video forensic analysis is much more difficult than the image analysis due to lossy video compression, so the current traces can be erased or significantly damaged by high compression rates, making all or part of the processing records unrecoverable. While numerous forensic methods have been developed based on digital images [3]–[9], the forensic analysis of videos has been less explored. It should be noted that methods based on images cannot also be applied

directly to videos [10]–[12]. This is due to some challenges such as compression, stabilization, scaling, and cropping, as well as the differences between frame types that can occur when producing a video.

Video identification algorithms are used to identify and distinguish camera types based on video produced by digital cameras. During the last few years, forensic specialists have been particularly interested in this topic. In general, there are two main ways to identify images and videos, namely, examining images or videos to extract a unique fingerprint of the camera and using metadata associated with the images or videos (the DNA of a video). Lopez *et al.* [13] demonstrated that the internal elements and metadata of video can be used for source video identification. Since metadata can be removed from an image or video, identifying video or images based on fingerprint is a reliable method. Moreover, two concepts are considered for identifying camera: individual source camera identification (ISCI) and source camera model identification (SCMI). ISCI distinguishes cameras from both the same and different camera models, while SCMI is a subset of ISCI that distinguishes a particular camera model from others, but cannot distinguish a particular camera model from the same camera models. SCMI has been more researched compared to ISCI [14]. Other aspect that is important in identifying source camera on video is codec for compensation purpose. The codec may affect the accuracy of the source camera identification, as some useful information is lost during the action.

As a result of the challenges and advances in research in the field of forensic video analysis, such as deep learning methods, there is a need for standard databases that allow researchers to more easily compare techniques using the same experimental protocols. Although there are several databases for identifying source cameras for images [15], [16], there are few databases for videos. Therefore, for new challenges such as ISCI and deep source camera identification analysis that focus on video, it is essential to have a database to perform new methods on video.

Since most video databases focus on videos recorded with a (videocassette recorder) VCR, and among them there is only one database for smartphones (Daxing) [1], in order to give new tasks to the databases, we focus on presenting a smartphone database for videos. It should be noted that the Daxing database cannot cover the ISCI challenge for all devices (out of the 22 models, 16 models can be used for the challenge) and QUFVD is more suitable to train a Deep Learning method due to the number of videos recorded in the database (6000 videos) compared to Daxing (1400). This study is an attempt to develop a database that is unique to the new challenges of smartphone video. The structure of the database for identifying source camera is shown in Figure 1. As shown in the figure, to evaluate the database, we need to extract frames. Generally, frames consists of intra-coded picture (I-frame), predictive coded picture (P-frame), and bi-predictive coded picture (B-frames), which show promising results among those obtained with I-frames [12], [14]. The database is

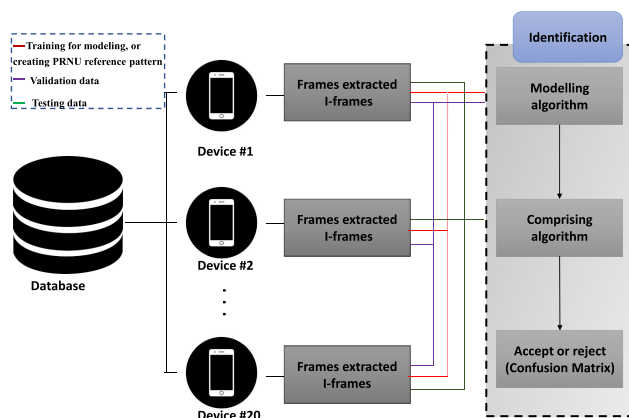


FIGURE 1. Overview of QUFVD structure and evaluation step for ISCI scenario.

presented by both videos and I-frames corresponding with videos. We also consider training, validation, and testing data for two common categories of methods used in the field, namely Photo Response Non Uniformity (PRNU) and Machine Learning approaches. PRNU, which is understood to be the unique fingerprint of the camera, is often referred to as residual noise or sensor pattern noise (SPN). PRNU occurs when the CCD (Charge Coupled Device) or CMOS (Complementary Metal Oxide Semiconductor) sensors process the input signal (light) and convert it into a digital signal. In deep learning methods, which are a popular category of machine learning, this training step should be performed to extract the fingerprint of the camera. The main challenges for these methods are the separation of content from noise and the number of training data. The first challenge can be solved by introducing architectures that address the problem by, for example, adding new layers and loss functions.

The paper is organized as follows. Section II is a review to available databases for videos with a brief description. Our motivation is explained in Section III. Our new video database is completely presented in Section IV. Section V describes database evaluation based on the deep learning method. The last section concludes this work.

II. LITERATURE REVIEW

The databases presented based on videos are summarized in Table 1.

One of the main reasons for the lower exploration of videos compared to images is that there are few standard digital video databases to develop the methods [14]. We explore the databases in the section.

CAMCOM2010 [17] is a contest designed to identify source YouTube videos. Two participants submitted results despite a satisfactory number of participants at first. However, the database is not available publicly.

The University of Surrey's website provides access to SULFA database [18]¹ that contains original videos and

¹<http://sulfa.cs.surrey.ac.uk/>

TABLE 1. The databases presented based on videos.

Method	Availability	Year	Description
SULFA [18]	Public	2012	150 videos collected from three camera sources
VISION [21]	Public	2017	34,427 images and 1914 videos, 35 portable devices of 11 major brands
video-ACID [14]	Public	2019	12 000 videos from 46 physical devices representing 36 unique camera models
Daxing [1]	Public	2019	43400 images and 1400 videos captured by 90 smartphones of 22 models belonging to 5 brands for capturing images
SOCRatES [24]	Public	2019	9700 images and 1000 videos captured with 103 different smartphones

forged ones. The original videos are suitable for source camera identification purpose. About 150 videos were collected from three sources that was also extended by [19]. Method presented in [20] was tested in the study.

VISION database² was introduced in [21] that is the most popular database in the field. In total, 35 portable devices from 11 major brands contributed 34,427 images and 1914 videos, all in native format and social format (Facebook, YouTube and WhatsApp are included). There are videos captured in indoor, outdoor, and flat scenarios. Videos of flat surfaces such as walls and sky are included in the flat scenario. Videos depicting offices or shops are included in the indoor scenario, while videos depicting gardens are included in the outdoor scenario. Three recording modes were used for each scenario: Still mode, where the user stands still while the video is recorded. While capturing the video (moving video), the user walks, the pan/tilt mode combines a pan with a rotation to achieve a recording. YouTube and WhatsApp social media platforms were used to exchange videos belonging to each scenario. In the study, they evaluated the database by method presented in [4].

video-ACID database³ was presented in [14] to source camera identification that is accessible publicly. Over 12,000 videos were collected from 46 physical cameras representing 36 different camera models in the video-ACID's database. All of these videos were shot manually to represent a range of lighting conditions, content, and motion. Moreover, this database is suitable for both SCMI or ISCI scenarios. They evaluated deep learning method presented in [22].

[1] presented a Daxing smartphone identification database,⁴ which include both images and videos from extensive smartphones of different brands, models and devices. The data from 90 smartphones, representing 22 models and 5 brands, includes 43400 images and 1400 videos. In the case of the iPhone 6S (Plus), 23 different smartphone models are available. Scenes selected normally include a sky, grass, rocks, trees, stairs, a vertical printer, a lobby wall, and a white wall in a classroom, among others. The videos were shot

vertically in each scene. Each scene contains at least three videos. In addition, all videos were recorded over 10 seconds. The database was evaluated by method presented in [23].

SOCRatES database⁵ [24] captured by smartphones. Around 9700 images and 1000 videos were taken by 103 different smartphones from 15 different brands. [3] and [25] were assessed on the database.

III. MOTIVATION

The rapid development of new smartphones in the field of imaging may be an important reason for the development of databases in forensic analysis, especially in the identification of source cameras. On the other hand, the coverage and completion of aspects that other databases have not considered in the development of the databases may lead researchers to present a new database.

As described in the previous sections, most databases contain videos recorded on VCR, and only one of these databases is dedicated to smartphones (Daxing) [1]. Although the database can be considered as an important database in this field as it covers a wide range of devices, there are some aspects that may lead researchers to develop a new database to meet new challenges in this field.

Table 2 shows the database in detail based on number of videos for each device. Of the 90 devices used in the database, 85 devices were considered for recording the videos. As can be seen from the table, the range of videos recorded by the devices is limited. This may be because the Daxing database focuses on both videos and images. The smallest number of videos recorded by a device in the Daxing database is 4 while the largest number of videos recorded by a device is 106 videos, where only one device has 106 videos and the rest of the devices have less than 31 videos. More devices have 12 to 28 videos. On average, the number of videos for each device is around 26. As a result, the assessment of PRNU-based methods may not be reliable. Furthermore, source camera identification techniques based on machine learning may face a problem of unbalanced data since the number of training videos is small and differs across the devices. This prompts the researchers to adjust and balance

²<https://lesc.dinfo.unifi.it/VISION/>

³misl.ece.drexel.edu/video-acid

⁴<https://github.com/xyhcn/Daxing>

⁵<http://socrates.eurecom.fr/>

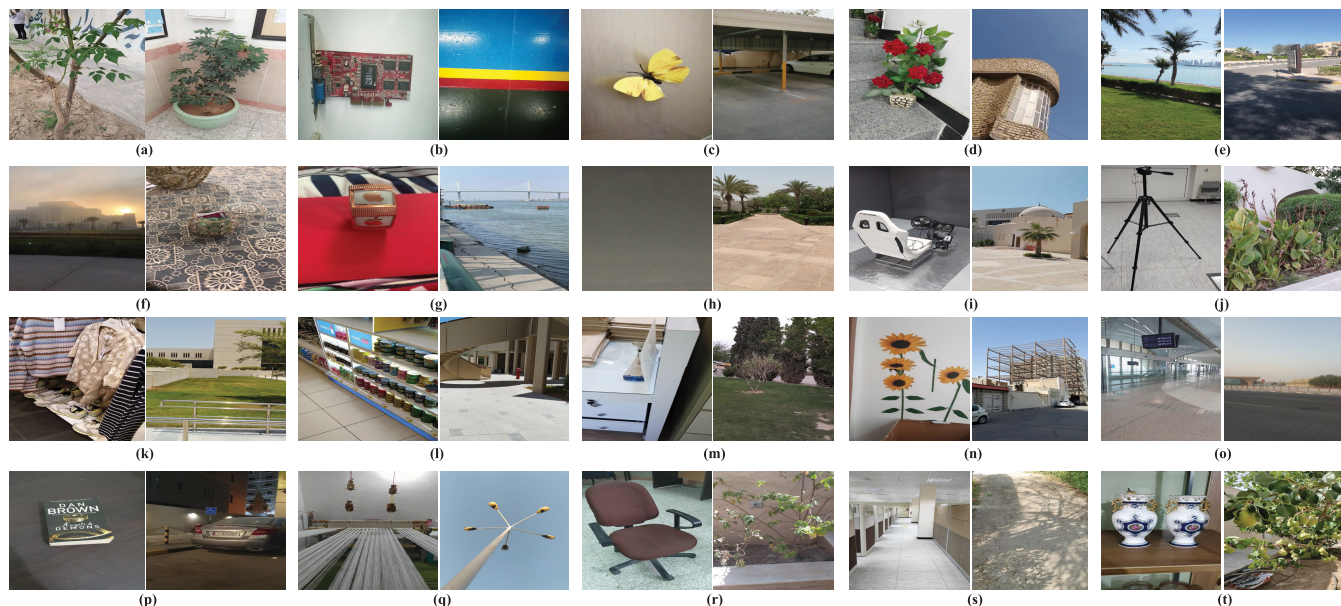


FIGURE 2. Sample frames from captured videos: (a) Huawei-Y7 (device 1), (b) Huawei-Y7 (device 2), (c) Huawei-Y9 (device1), (d) Huawei-Y9 (device 2), (e) iPhone-8Plus (device 1), (f) iPhone-8Plus (device 2), (g) iPhone-XsMax (device 1), (h) iPhone-XsMax (device 2), (i) Nokia-5.4 (device 1), (j) Nokia-5.4 (device 2), (k) Nokia-7.1 (device 1), (l) Nokia-7.1 (device 2), (m) Samsung-A50 (device 1), (n) Samsung-A50 (device 2), (o) Samsung-Note9 (device 1), (p) Samsung-Note9 (device 2), (q) Xioami-RedmiNote8 (device 1), (r) Xioami-RedmiNote8 (device 2), (s) Xioami-RedmiNote9Pro (device 1), (t) Xioami-RedmiNote9Pro (device 2).

the database before use. For example, for the iPhone 8 Plus, 24 videos were recorded for device #1 and only 4 videos were recorded for device #2.

As our experiments have shown (Section V), increasing the number of training data can improve the results in our database. Also, since most machine learning methods require enough data to train, it is obvious that a database with many more videos is better suited to machine learning methods compared to Daxing for the ISCI scenario. As shown in Table 2, for implementing ISCI scenario, only one device has 106 videos to train, and the rest of the devices have less than 31 videos to train. Additionally, to design the structure of the Daxing database to be suitable for a machine learning approach, the videos need to be divided into training, testing and validation sets. If we consider 26 videos for each device as the average of the database and use our structure to split the database, we have 15, 7 and 4 videos for training, testing and validation respectively. Therefore, it is obvious that it can be less to compare the machine learning methods fairly. It should be noted that the Daxing database may be more suitable for machine learning methods for the SCMI scenario for more models.

Finally, it should be noted that a new database can be connected to other databases such as Daxing to obtain more data and deal with new challenges.

IV. QUFVD DESCRIPTION

In this section, we discuss the features and structure of QUFVD. For describing a database, the following options are important: number of videos and camera, resolution, codec,

suitable for SCMI or ISCI. The options are described in more detail in this subsection. Table 3 summarizes our database with its features. The QUFVD is publicly available.⁶

A. DEVICES

There are several popular manufacturers that produce different smartphone brands. Among them, there are few brands that are widely used by people. In order to have a variety of brands, we collected the devices to be used for video recording and selected 5 popular brands: iPhone, Samsung, Huawei, Xioami and Nokia. For each brand, we selected two different models, and for each model, we selected two devices. Therefore, four devices are considered for each brand. The total number of devices used to collect this database should be 20 devices.

B. SIZE PROPERTIES

With the development of deep learning methods in this area, a large number of videos or frame can improve the results in this area, as shown in this article. Therefore, a database with suitable size can be considered for both traditional methods such as PRNU and deep learning methods. In our database, 300 videos are collected for each device, making a total of 6000 videos available. The length of the videos is between 11 and 15 seconds at a frame rate of 30 frames per second. Since I-frames play an important role in identifying the source [12], [14], these types of frames are also extracted. Depending on the length and content of each video, they are

⁶<https://www.dropbox.com/sh/nb543na9qq0wlaz/AAAc5N8ecjawk2K1VF8kfkrya?dl=0>

TABLE 2. Number of videos captured based on Daxing database for each device.

Smartphone	Number of videos for each device (separated by comma)	Smartphone	Number of videos for each device (separated by comma)
Huawei P20 (5 devices)	26, 24, 24, 30, 17	iPhone 7 Plus (5 devices)	20, 22, 21, 24, 24
Huawei Mate 9 (3 devices)	29, 21, 25	iPhone 8 Plus (2 devices)	24, 4
Huawei Mate 9 Plus (1 device)	24	OPPO R9 (6 devices)	24, 24, 24, 24, 24, 20
Huawei P9 (5 devices)	24, 24, 24, 24, 25	OPPO R9S Plus (1 device)	23
Huawei P10 (3 devices)	22, 16, 24	OPPO R11 (5 devices)	18, 24, 26, 25, 19
Huawei 10 Plus (4 devices)	24, 24, 23, 24	OPPO R 11T (1 device)	24
iPhone 9 (5 devices)	21, 24, 22, 24, 27	VIVO X9 (4 devices)	21, 25, 24, 24
iPhone 6 Plus (4 devices)	106, 24, 24, 22	VIVO X9 Plus (1 devices)	24
iPhone 6S (12 devices)	24, 23, 24, 26, 24, 25, 25, 27, 24, 24, 24, 24	VIVO X9I (1 devices)	8
iPhone 6S plus (9 devices)	18, 25, 24, 22, 24, 24, 24, 24, 22	VIVO Y85 (1 device)	12
iPhone 7 (2 devices)	24, 27	Xiaomi 4A (5 devices)	12, 12, 12, 12, 12

TABLE 3. The devices of our database with their characteristics.

Brand	Model	Resolution	Number of videos	Number of I-frame	Length in Secs	Operating system
Samsung	Galaxy A50 (device #1)	1080 × 1920	300	3654	11-15	Android 10.0
Samsung	Galaxy A50 (device #2)	1080 × 1920	300	3782	11-15	Android 10.0
Samsung	Note9 (device #1)	1080 × 1920	300	3956	11-15	Android 10.0
Samsung	Note9 (device #2)	1080 × 1920	300	3962	12-15	Android 10.0
Huawei	Y7 (device #1)	720 × 1280	300	3630	11-15	Android 9.0
Huawei	Y7 (device #2)	720 × 1280	300	3642	11-15	Android 9.0
Huawei	Y9 (device #1)	720 × 1280	300	4146	11-14	Android 9.0
Huawei	Y9 (device #2)	720 × 1280	300	4011	11-15	Android 9.0
iPhone	8 Plus (device #1)	1080 × 1920	300	3991	11-15	iOS 13
iPhone	8 Plus (device #2)	1080 × 1920	300	4080	11-14	iOS 13
iPhone	XS Max (device #1)	1080 × 1920	300	3893	11-15	iOS 13
iPhone	XS Max (device #2)	1080 × 1920	300	4074	11-15	iOS 13
Nokia	5.4 (device #1)	1080 × 1920	300	3350	11-13	Android 10.0
Nokia	5.4 (device #2)	1080 × 1920	300	3531	11-14	Android 10.0
Nokia	7.1 (device #1)	1080 × 1920	300	3904	11-13	Android 10.0
Nokia	7.1 (device #2)	1080 × 1920	300	3819	11-14	Android 10.0
Xiaomi	Redmi Note8 (device #1)	1080 × 1920	300	3776	11-14	Android 11.0
Xiaomi	Redmi Note8 (device #2)	1080 × 1920	300	3598	11	Android 11.0
Xiaomi	Redmi Note9 Pro (device #1)	1080 × 1920	300	3888	11-15	Android 11.0
Xiaomi	Redmi Note9 Pro (device #2)	1080 × 1920	300	3838	11-13	Android 11.0

different. A total of 76531 I-frames are extracted by FFmpeg⁷ software. Finally, to test Deep Learning methods, 500 patches of 350×350 are extracted from each I-frame. A total of 980,580 Patches are extracted for train set.

C. CONTENT PROPERTIES

In this database collection, we rely mainly on the static camera despite the static and moving state of the camera when recording. This database contains very diverse video collections of different scenes, either outdoor, indoor, moving or still objects, mainly gardens, sky, streets, shops, domestic staff and sea. Figure 2 shows samples of the data based on each device.

D. ISCI PROPERTIES

One way to challenge the identification of the source camera is that the captured videos are from smartphones with the

same camera model. In the database, two devices are considered for each smartphone. For example, for Samsung Galaxy A50, the videos were captured from two devices. Moreover, the challenge is studied in the evaluation section. Therefore, our database contains both SCMI and ISCI scenarios for all models defining two 10- and 20-class problems.

E. CODEC PROPERTIES

Video files are compressed with codecs, which are always a tradeoff between quality and size (better quality vs. larger file size). Video files can be compressed to reduce their size, which can reduce bandwidth usage and increase streaming speed. For encoding high-definition video, AVC is the standard codec used by several online video services, including YouTube and Vimeo. The MPEG-4 and the H.264 standards are implemented by the library 'libx264' in FFmpeg. All smartphones used for our database recorded videos according to the H.264 video encoding standard, except for the iPhone Xs Max and the Samsung Note9 (H.265).

⁷<https://www.ffmpeg.org/>

F. I-FRAME PROPERTIES

Group of Pictures (GOP) consists of I-frames, P-frames and B-frames as intra-coded picture, predictive-coded picture and bi-predictive coded picture respectively in coding standards like MPEG series and H.264. I-frames are the least compressible and do not require other video frames for decoding. P-frames can be decompressed with data from previous frames and are more compressible than I-frames. For B-frames, previous and forward frames can also be used as data references to achieve the highest compression. The I-frame is generally more detailed than the P- and B-frames. The GOP size, generally divided into fixed and unfixed, is the number of B- and P-frames between two consecutive I-frames. Several studies have demonstrated that methods based on I-frames give better results compared to other frames [26]–[28].

G. VIDEO NAMING

All videos are renamed according to the following rule: “Brand_Model_device No._Video No.” for example: We have “iPhone_XS_Max_1_(1)”, which refers to the first video from the first device of the model XS Max of the iPhone mobile brand, another example: “Samsung_A50_2_(30)”, refers to the video number 30 from the second device of the model Galaxy A50 of the mobile brand Samsung, and so on. Also, for I-frames, the frame number and type are appended to the video name, e.g., “Huawei-Y7Prime2019-1(14)-31-I” indicates the 31st frame with the type I-frame of the device.

H. RESOLUTION AND COLOR MODE

The resolution of a video is the width and height of the video in pixels. All videos recorded in the database are based on the rear camera of smartphones. There are two types of resolutions in the database, namely 720×1280 and 1080×1920 . Also, the frames are stored in two modes: Color (True Color) and Grayscale. It can be tested whether the resolution and color mode affect the results.

I. STRUCTURE OF THE DATABASE

The overall structure of QUFVD is shown in Figure 3. The structure can be modified by researchers according to their methods and facilities. Moreover, the database can be combined with other databases (e.g. the Daxing database) to address new scenarios and challenges and to have a wider choice of brands.

V. QUFVD EVALUATION

The quality of our database is evaluated in this section by experimenting with ISCI and SCMI scenarios with different settings based on a Deep Learning method. We divide the experiments into different scenarios showing the influence of some conditions on the results. This result provides a baseline for the accuracy of camera model identification in the QUFVD database and can be used for comparison with other methods. The division of the database for the experiments is

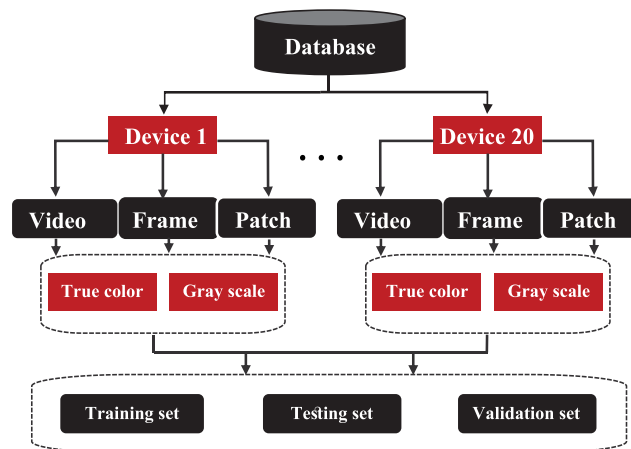


FIGURE 3. Structure of the folders in the database.

that 80% of these videos are considered as training set and the remaining 20% are considered as test set. Also, 20% of the training data is considered as validation data. The structure can be used for both classical (like PRNU) and machine learning methods. For example, in PRNU methods, reference patterns can be obtained from videos in the training data and query patterns in the test data. Since, as mentioned earlier, the I-frames lead to better results, the I-frames of the videos are extracted to evaluate the database. The statistics for training, testing, and validation at both the video and frame levels are shown in Table 4. For each video in each experimental series, we selected all I-frames related to the videos in the training, testing, and validation series. A total of 76531 I-frames were extracted.

The method presented in [29] is used to evaluate our database. Also, in references [30], [31] and [22], the CNN method (MISLnet CNN architecture [29]) was used to identify the source camera, using frames to train the network. The network used a constrained convolutional layer that was added as the first layer that used three kernels with size 5. This layer is constructed in such a way that there are relationships between adjacent pixels that are independent of the content of the scene. The methods was tested on VISION database [21]. The experiments showed that the layer can improve results compared with deep learning architectures without the layer. The structure of the CNN for the three studies is shown in Figure 4. As shown in the figure, a constrained convolutional layer is added to a simple CNN.

Our database is evaluated against two main scenarios of ISCI and SCMI. A 10-class problem should be considered for SCMI and a 20-class problem for ISCI. For each, the effect of the number and size of patches is examined, as well as the effect of the color mode, i.e., gray and true color modes. All videos were encoded according to the respective device codec using the H.264 or H.265 video encoding standard, and no video was edited or re-encoded.

To identify a video based on its I-frames, all I-frames in the test set are considered. The scores obtained by the CNN

based on the highest probability show which I-frames belong to which classes. At the video level, a majority vote then decides all the frames that belong to a video.

A 64-bit operating system (Ubuntu 18) with a CPU E5-2650 v4 @ 2.20 GHz, 128.0 GB RAM, and four NVIDIA GTX TITAN X. was used in order to run our experiments.

A. ISCI VS SCMI

The performance of the network is measured by computing the accuracy based on frame-level and video-level in both scenarios ISCI and SCMI. In classification stage, each frame/video in the test data is classified into one of the 10-class (SCMI) or 20-class (ISCI). The frame-level and video-level results for SCMI scenarios for each smartphone model are shown in Table 5.

To investigate the effects of device dependency, the ISCI scenario is considered. The results of the frame and video levels in terms of accuracy for the ISCI scenario for each device are shown in Table 6. The result is based on the accuracy for each device.

Also overall accuracy, precision, recall and F1-score based on frame-level for both scenarios ISCI and SCMI are reported in Table 7. Precision is also called Positive Predictive Value (PPV) which is a measure of the closeness of the set of predicted results. Recall is also known as True Positive Rate (TPR) and F1-score is the harmonic average of the precision and recall, where it is at its best at a value of 1 meaning perfect precision and recall.

Tables 5 and 6 also list the effects of color mode, i.e., grayscale and true color. With this premise, Figure 5 and 6 provide a more comprehensive picture of camera identification performance to check the quality of the CNN by presenting the Receiver Operating Characteristic (ROC) curves for a selected group of ten and twenty cameras from our database. Two values are calculated for each threshold: True Positive Ratio (TPR) and False Positive Ratio (FPR). The TPR of a given class, e.g. Huawei Y7, is the number of outputs whose actual and predicted class is Huawei Y7 divided by the number of outputs whose predicted class is Huawei Y7. The FPR is calculated by dividing the number of outputs whose actual class is not Huawei Y7, but whose predicted class was Huawei Y7 by the number of outputs whose predicted class is not Huawei Y7.

One of the most important factors in machine learning is how much training data the model needs to perform well. To show how it works, a series of experiments were conducted with increasing training data for the SCMI scenario for both gray and color modes. Table 8 shows the effect of the factor.

In addition, the size of the patches can affect the performance of the CNN methods. For the experiment, four different sizes were considered for the SCMI scenario based on 10000 greyscale patches (see Table 9).

For a more detailed analysis of the error detections, the confusion matrix for the ISCI scenario in grayscale mode is given in Table 10.

Also, the processing time for patch, frame and a video with 11 I-frames is shown in Table 11, which is performed by processing frame size of 1920×1080 .

B. RESULT DISCUSSION

State-of-the-art source camera identification methods have faced challenges such as compression, stabilization, and ISCI. Various methods have been presented to overcome these challenges. Recently, Deep Learning methods have been introduced to solve these challenges. As mentioned earlier, our database is also evaluated using deep learning method developed to solve these problems. Overall, the results in frame and video levels show that the method is successful for the SCMI problem, but it does not work well for the ISCI challenge. For both scenarios, when the results are reported at the video level, improvement can be seen. The results are discussed in more detail below.

As shown in Table 5 in frame level, all devices except the Y7, 8 Plus, and Redmi Note9 Pro achieve more than 70 % accuracy in grayscale mode. The biggest improvement in the mode is for the Note 9 compared to the color mode. The best results are reported for the Note 9 and Xs Max, which have the same codec (H.265). At the video level, an overall improvement is seen for all devices. The best result with 95% is also obtained for the Note 9. However, its codec is similar to that of Xs Max, so Xs Max cannot see an improvement like the Note 9. In general, we cannot make the decision that the codec has a direct effect on the results. Moreover, we can see that the resolution does not affect the results, since Y7 and Y9 have the lowest resolution, but their results are not worse. Therefore, as a result of this work, the two cases cannot confirm that codec and resolution are two effective factors in this area. However, in grayscale mode, it can be seen that the results are better than in color mode.

Based on Table 6 (ISCI scenario), although only 3 devices have an accuracy of less than 65%, half of the devices achieve an accuracy of less than 50% at the frame level. Even though Note 9 for Device 1 scores the best among all devices, similar to the SCMI scenario, Device 2 scores only 66.7%, which means the fifth place. There are no meaningful results in the table, except that grayscale mode still performs better than color mode for both frame and video levels.

Figures 5 and 6 show the TPR compared to the FPR for the two scenarios SCMI and ISCI in two modes (color and grayscale) at different frame-level thresholds. As can be seen from the figures, we have different analysis for the devices in terms of TPR and FPR. The best performance is shown by Nokia 5.4 with Area Under Curve (AUC=0.989) compared to the second ranked Note 9 with AUC=0.987 in grayscale mode (Figure 5 (b)). Moreover, as shown in Figure 6 (a and b), RedmiNote9Pro performs significantly better in grayscale mode. From the figure, it can be seen that Note 9 device 1 has the best performance with AUC=0.989.

As shown in Table 7, all metrics are better in the SCMI scenario than in the ISCI scenario. Based on the results,

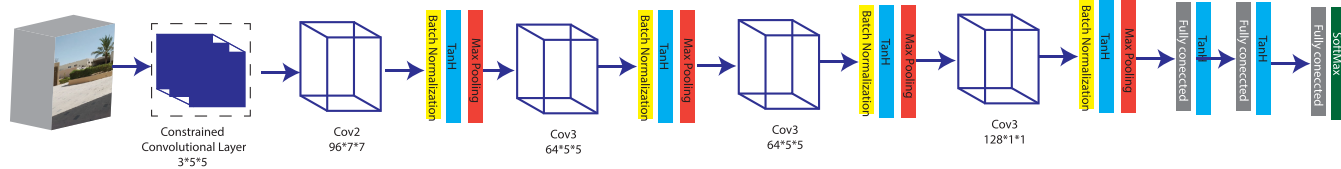


FIGURE 4. Architecture of the ConstrainedNet. (Based on [29]).

TABLE 4. The statistics for training, testing, and validation at both the video and frame levels.

Brand	Model	Video			I-frame		
		Training	Testing	Validation	Training	Testing	Validation
Samsung	Galaxy A50 (device #1)	192	60	48	2335	726	593
Samsung	Galaxy A50 (device #2)	192	60	48	2476	727	579
Samsung	Note9 (device #1)	192	60	48	2546	788	628
Samsung	Note9 (device #2)	192	60	48	2534	790	638
Huawei	Y7 (device #1)	192	60	48	2307	723	600
Huawei	Y7 (device #2)	192	60	48	2342	724	576
Huawei	Y9 (device #1)	192	60	48	2662	826	658
Huawei	Y9 (device #2)	192	60	48	2562	815	634
iPhone	8 Plus (device #1)	192	60	48	2534	827	630
iPhone	8 Plus (device #2)	192	60	48	2597	816	667
iPhone	XS Max (device #1)	192	60	48	2469	788	636
iPhone	XS Max (device #2)	192	60	48	2626	806	642
Nokia	5.4 (device #1)	192	60	48	2142	671	537
Nokia	5.4 (device #2)	192	60	48	2272	702	557
Nokia	7.1 (device #1)	192	60	48	2502	787	615
Nokia	7.1 (device #2)	192	60	48	2447	763	609
Xiaomi	Redmi Note8 (device #1)	192	60	48	2420	764	592
Xiaomi	Redmi Note8 (device #2)	192	60	48	2304	721	573
Xiaomi	Redmi Note9 Pro (device #1)	192	60	48	2492	778	618
Xiaomi	Redmi Note9 Pro (device #2)	192	60	48	2460	769	609
Total		3840	1200	960	49029	15311	12191

TABLE 5. The results of the frame and video levels in terms of accuracy (%) for the SCMI scenario for each smartphone model.

Model	I-frame		Video	
	Gray	Color	Gray	Color
Galaxy A50	72.8	68.8	73.3	70.0
Note9	78.7	72.0	95.8	87.6
Y7	68.0	70.4	84.2	75.8
Y9	76.9	76.9	86.7	90.8
8 Plus	67.8	69.6	84.2	68.3
XS Max	76.8	69.3	68.3	70.0
5.4	81.8	81.6	90.8	91.7
7.1	75.5	76.7	90.0	79.8
Redmi Note8	75.8	72.8	80.8	84.2
Redmi Note9 Pro	66.4	65.0	65.8	70.0
Overall accuracy	74.0	72.5	82.0	78.8

it is essential to improve the scenario in machine learning approaches.

Table 8 shows that increasing the number of patches to be trained can improve the results in both gray scale and color modes for SCMI at frame level. 37.3% is improved from 5000 patches to all patches (about 90000) trained for each class. In the ISCI scenario, the improvement is 49.9% if all patches (about 45000) are trained, which is an improvement of about 6%. Also in gray mode, when all

TABLE 6. The results of the frame and video levels in terms of accuracy (%) for the ISCI scenario for each device.

Model	I-frame		Video	
	Gray	Color	Gray	Color
Galaxy A50 (#1)	35.8	36.9	48.3	56.7
Galaxy A50 (#2)	29.3	29.2	46.7	25.0
Note9 (#1)	67.1	57.6	95.0	80.0
Note9 (#2)	58.8	44.4	66.7	63.3
Y7 (#1)	44.1	47.8	53.3	45.0
Y7 (#2)	51.5	52.2	68.3	75.0
Y9 (#1)	79.2	78.2	65.0	83.3
Y9 (#2)	59.4	62.4	63.3	66.7
8 Plus (#1)	35.6	47.5	60.0	36.7
8 Plus (#2)	30.0	25.3	25.0	20.0
XS Max (#1)	59.2	61.8	60.0	53.3
XS Max (#2)	47.6	47.8	71.7	53.3
5.4 (#1)	39.2	30.8	58.3	31.7
5.4 (#2)	35.7	37.5	28.3	35.0
7.1 (#1)	54.9	43.8	61.7	53.3
7.1 (#2)	48.1	41.3	46.7	40.0
Redmi Note8 (#1)	70.8	51.8	78.3	50.0
Redmi Note8 (#2)	53.5	60.4	88.3	63.3
Redmi Note9 Pro (#1)	45.9	51.2	63.3	48.3
Redmi Note9 Pro (#2)	51.1	39.8	43.3	43.3
Overall accuracy	49.9	47.9	59.6	51.2

patches are trained, the results are 2 % higher than in color mode.

TABLE 7. The results of the frame level in terms of accuracy, PPV, TPV and F-score (%) for the SCMI and ISCI scenario.

	Accuracy	PPV	TPR	F1-score
SCMI (Gray)	74.00	74.04	74.01	74.03
ISCI (Gray)	49.90	49.84	49.76	49.80

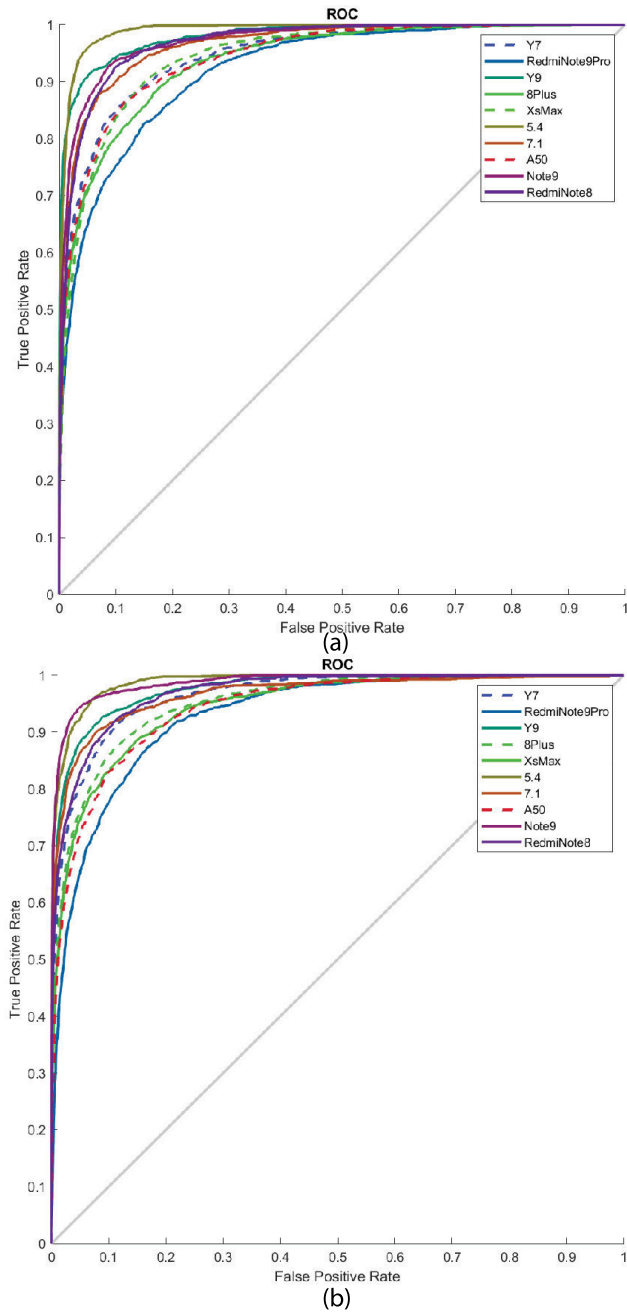


FIGURE 5. True and false positive rates (ROC) obtained in SCMI scenario (a) 10 classes with color mode (b) 10 classes with gray scale mode.

Table 9 shows that while the size of each patch can improve performance, it is limited by the size of the 350×350 . It should be noted that the experiment was conducted

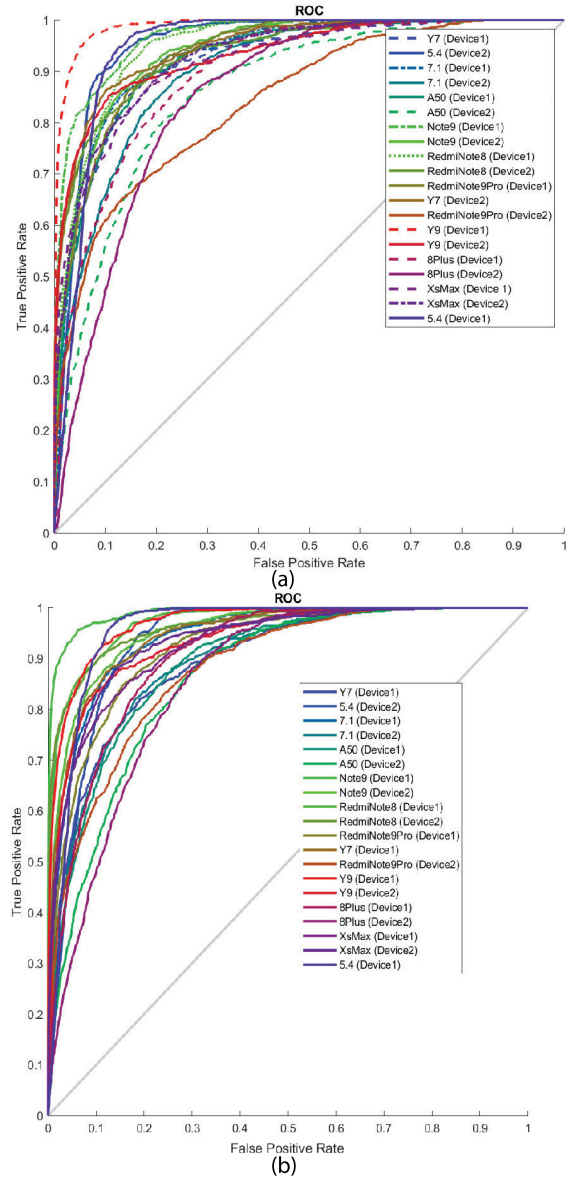


FIGURE 6. True and false positive rates (ROC) obtained in ISCI scenario (a) 20 classes with color mode (b) 20 classes with gray scale mode.

TABLE 8. The impact of training data set on performance.

	5000	10000	30000	60000	All patches
SCMI (Gray)	41.5	54.1	63.4	69.8	74.0
SCMI (Color)	35.4	55.3	63.0	69.5	72.5
ISCI (Gray)	43.4	44.0	47.5	-	49.9
ISCI (Color)	41.2	43.2	46.4	-	47.9

TABLE 9. The impact of patch size on performance.

	128×128	256×256	350×350	512×512
SCMI (Gray)	40.8	51.0	54.1	52.8

with 10000 patches for each class. For sizes over 350×350 , a drop in performance is indicated. Therefore, we chose this size for all experiments in the evaluation.

TABLE 10. Confusion matrix of ISCI scenario in grayscale mode. Classes 1 to 20 are Y7 (device 1), Nokia 5.4 (device 2), Nokia 7.1 (device 1), Nokia 7.1 (device 2), A50 (device 1), A50 (device 2), Note 9 (device 1), Note 9 (device 2), RedmiNote8 (device 1), RedmiNote8 (device 2), RedmiNote9Pro (device 1), Y7 (device 2), RedmiNote9Pro (device 2), Y9 (device 1), Y9 (device 2), 8 Plus (device 1), 8 Plus (device 2), Xs Max (device 1), Xs Max (device 2), Nokia 5.4 (device 1), respectively.

303	2	4	42	26	20	1	1	11	4	33	64	16	32	23	33	22	25	18	7	44.1 (%)
14	174	14	11	4	6	0	2	17	0	17	1	22	1	8	0	0	1	0	196	35.7 (%)
11	25	375	142	16	13	0	12	10	17	9	8	14	4	0	0	5	1	9	12	54.9 (%)
19	17	136	284	12	1	3	9	7	11	15	8	13	0	10	8	6	6	11	15	48.1 (%)
36	18	14	23	291	193	1	12	10	9	22	17	57	7	21	7	17	30	0	27	35.8 (%)
38	23	18	12	159	232	3	21	23	7	52	5	66	4	24	29	11	34	5	26	29.3 (%)
2	11	2	3	8	7	703	133	0	9	17	1	4	1	20	29	28	17	51	2	67.1 (%)
14	6	1	20	47	30	23	487	0	10	0	2	5	35	31	35	17	24	29	12	58.8 (%)
10	17	10	2	13	31	0	1	529	16	18	22	26	11	8	4	3	12	1	13	70.8 (%)
36	5	96	60	8	8	4	5	56	573	54	79	44	3	4	6	5	8	9	9	53.5 (%)
14	13	45	67	10	48	6	3	15	16	385	12	89	12	17	7	5	12	38	24	45.9 (%)
97	8	33	46	15	23	7	3	25	8	14	437	33	10	26	23	17	15	2	6	51.5 (%)
34	12	8	23	13	16	2	3	24	3	25	15	242	5	12	1	5	11	8	12	51.1 (%)
5	2	1	1	1	2	0	3	3	0	0	1	12	457	79	3	3	2	2	0	79.2 (%)
1	1	0	0	8	24	0	6	1	2	5	0	11	165	431	4	44	20	0	2	59.4 (%)
9	9	6	7	34	35	12	51	3	9	12	8	30	16	35	407	355	45	60	1	35.6 (%)
12	2	3	0	10	4	1	15	6	4	17	14	15	23	28	116	154	65	25	0	30.0 (%)
20	0	2	3	11	15	0	4	1	13	5	12	21	15	23	23	45	388	52	2	59.2 (%)
27	3	8	14	34	11	22	19	5	8	69	17	26	24	6	92	73	72	486	4	47.6 (%)
21	354	11	3	6	8	0	0	18	2	9	1	23	1	9	0	1	0	0	301	39.2 (%)
41.9 (%)	24.8 (%)	47.6 (%)	37.2 (%)	40.1 (%)	31.9 (%)	89.2 (%)	61.6 (%)	69.2 (%)	79.5 (%)	49.5 (%)	60.4 (%)	31.5 (%)	55.3 (%)	52.9 (%)	49.2 (%)	18.9 (%)	49.2 (%)	60.3 (%)	44.9 (%)	49.9 (%)

TABLE 11. The processing time (second) for patch, frame and a video with 11 I-frames.

128 × 128			256 × 256			350 × 350			512 × 512		
Patch	Frame	Video	Patch	Frame	Video	Patch	Frame	Video	Patch	Frame	Video
0.491	0.723	1.500	0.517	0.905	2.812	0.616	0.993	3.300	0.817	1.330	3.920

Table 10 shows the confusion matrix obtained for the ISCI scenario in grey scale mode. As mentioned earlier, the scenario is more challenging than SCMI and the results can be improved in the next studies. The confusion matrix can show misclassifications between all classes. As shown in the table, misclassifications between devices of the same brand occur in most cases, e.g., classes 14 and 15 (two devices Y9) have the most misclassifications when they misidentify each other.

As can be seen in Table 11, the time increases in all patch, frame and video levels as the patch size increases.

VI. CONCLUSION

This paper presents a new video database (QUFVD) based on smartphones for source camera identification. The database includes five popular smartphone brands with two models per brand with two devices for each model, 6000 original videos, and 76531 I-frames. The entire database is provided with an evaluation analysis for use by the research community.

The database is suitable for new challenges such as ISCI and for use by deep learning methods. The results show that improvement is essential for ISCI. Although it is not a fair comparison, the Deep Learning method used in our study achieves promising results compared to the results reported by Daxing, which are based on the PRNU method.

In order to improve the video level results, different decision making approaches such as fusion methods based on weighting the score of the classifiers can be applied in the future. We will add a few more tasks to the database where we transfer videos over social media such as WhatsApp and Facebook to study the impact of compression on source camera identification. To detect video tampering, another task adds forged videos to the database. To get more data and new challenges, our database can be attached to other databases. A augmentation method can be applied on the database to have more data to train them. Although we cannot clearly see the effects of codec and resolution in the method, it can be studied by other methods.

ACKNOWLEDGMENT

This publication was made possible by NPRP grant # NPRP12S-0312-190332 from Qatar National Research Fund (a member of Qatar Foundation). Open Access funding provided by the Qatar National Library. The statement made herein are solely the responsibility of the authors.

REFERENCES

- [1] H. Tian, Y. Xiao, G. Cao, Y. Zhang, Z. Xu, and Y. Zhao, "Daxing smartphone identification dataset," *IEEE Access*, vol. 7, pp. 101046–101053, 2019.
- [2] S. Milani, M. Fontani, P. Bestagini, M. Barni, A. Piva, M. Tagliasacchi, and S. Tubaro, "An overview on video forensics," *APSIPA Trans. Signal Inf. Process.*, vol. 1, no. 1, p. e2, 2012, doi: [10.1017/ATSIP.2012.2](https://doi.org/10.1017/ATSIP.2012.2).
- [3] J. Lukáš, J. Fridrich, and M. Goljan, "Digital camera identification from sensor pattern noise," *IEEE Trans. Inf. Forensics Security*, vol. 1, no. 2, pp. 205–214, Jun. 2006.
- [4] M. Chen, J. Fridrich, M. Goljan, and J. Lukáš, "Determining image origin and integrity using sensor noise," *IEEE Trans. Inf. Forensics Security*, vol. 3, no. 1, pp. 74–90, Mar. 2008.
- [5] A. Lawgaly and F. Khelifi, "Sensor pattern noise estimation based on improved locally adaptive DCT filtering and weighted averaging for source camera identification and verification," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 2, pp. 392–404, Feb. 2017.
- [6] A. Lawgaly, F. Khelifi, and A. Bouridane, "Weighted averaging-based sensor pattern noise estimation for source camera identification," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2014, pp. 5357–5361.
- [7] X. Kang, Y. Li, Z. Qu, and J. Huang, "Enhancing source camera identification performance with a camera reference phase sensor pattern noise," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 2, pp. 393–402, Apr. 2012.
- [8] F. Ahmed, F. Khelifi, A. Lawgaly, and A. Bouridane, "Comparative analysis of a deep convolutional neural network for source camera identification," in *Proc. IEEE 12th Int. Conf. Global Secur., Saf. Sustainability (ICGS3)*, Jan. 2019, pp. 1–6.
- [9] F. Ahmed, F. Khelifi, A. Lawgaly, and A. Bouridane, "Temporal image forensic analysis for picture dating with deep learning," in *Proc. Int. Conf. Comput., Electron. Commun. Eng. (iCCECE)*, Aug. 2020, pp. 109–114.
- [10] M. Iuliani, M. Fontani, D. Shullani, and A. Piva, "Hybrid reference-based video source identification," *Sensors*, vol. 19, no. 3, p. 649, 2019.
- [11] S. Mandelli, P. Bestagini, L. Verdoliva, and S. Tubaro, "Facing device attribution problem for stabilized video sequences," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 14–27, 2019.
- [12] E. Altinisik and H. T. Sencar, "Source camera verification for strongly stabilized videos," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 643–657, 2021.
- [13] R. R. López, E. A. Luengo, A. L. S. Orozco, and L. J. G. Villalba, "Digital video source identification based on container's structure analysis," *IEEE Access*, vol. 8, pp. 36363–36375, 2020.
- [14] B. C. Hosler, X. Zhao, O. Mayer, C. Chen, J. A. Shackelford, and M. C. Stamm, "The video authentication and camera identification database: A new database for video forensics," *IEEE Access*, vol. 7, pp. 76937–76948, 2019.
- [15] T. Gloe and R. Böhme, "The 'Dresden image database' for benchmarking digital image forensics," in *Proc. ACM Symp. Appl. Comput.*, 2010, pp. 1584–1590.
- [16] 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.
- [17] W. van Houten, Z. Geradts, K. Franke, and C. Veenman, "Verification of video source camera competition (CAMCOM 2010)," in *Proc. Int. Conf. Pattern Recognit.* Berlin, Germany: Springer, 2010, pp. 22–28.
- [18] G. Qadir, S. Yahaya, and A. T. S. Ho, "Surrey university library for forensic analysis (SULFA) of video content," in *Proc. IET Conf. Image Process. (IPR)*. Edison, NJ, USA: IET, 2012, pp. 1–6.
- [19] L. D'Amiano, D. Cozzolino, G. Poggi, and L. Verdoliva, "Video forgery detection and localization based on 3D patchmatch," in *Proc. IEEE Int. Conf. Multimedia Expo Workshops (ICMEW)*, Jun. 2015, pp. 1–6.
- [20] K. Rosenfeld and H. T. Sencar, "A study of the robustness of PRNU-based camera identification," *Proc. SPIE*, vol. 7254, Feb. 2009, Art. no. 72540M.
- [21] 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, pp. 1–16, Dec. 2017.
- [22] B. Hosler, O. Mayer, B. Bayar, X. Zhao, C. Chen, J. A. Shackelford, and M. C. Stamm, "A video camera model identification system using deep learning and fusion," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 8271–8275.
- [23] M. Goljan, J. Fridrich, and T. Filler, "Large scale test of sensor fingerprint camera identification," *Proc. SPIE*, vol. 7254, Feb. 2009, Art. no. 72540I.
- [24] C. Galdi, F. Hartung, and J.-L. Dugelay, "SOCRaTES: A database of realistic data for SOURCE camera REcognition on smartphones," in *Proc. 8th Int. Conf. Pattern Recognit. Appl. Methods*, 2019, pp. 648–655.
- [25] C.-T. Li, "Source camera identification using enhanced sensor pattern noise," *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 2, pp. 280–287, Jun. 2010.
- [26] T. Höglund, P. Brolund, and K. Norell, "Identifying camcorders using noise patterns from video clips recorded with image stabilisation," in *Proc. 7th Int. Symp. Image Signal Process. Anal. (ISPA)*, Sep. 2011, pp. 668–671.
- [27] M. Chen, J. Fridrich, M. Goljan, and J. Lukáš, "Source digital camcorder identification using sensor photo response non-uniformity," *Proc. SPIE*, vol. 6505, Mar. 2007, Art. no. 65051G.
- [28] M. Goljan, M. Chen, P. Comesaña, and J. Fridrich, "Effect of compression on sensor-fingerprint based camera identification," *Electron. Imag.*, vol. 2016, no. 8, pp. 1–10, Feb. 2016.
- [29] B. Bayar and M. C. Stamm, "Constrained convolutional neural networks: A new approach towards general purpose image manipulation detection," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 11, pp. 2691–2706, Nov. 2018.
- [30] D. Timmerman, S. Bennabhaktula, E. Alegre, and G. Azzopardi, "Video camera identification from sensor pattern noise with a constrained ConvNet," 2020, *arXiv:2012.06277*.
- [31] O. Mayer, B. Hosler, and M. C. Stamm, "Open set video camera model verification," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2020, pp. 2962–2966.



YOUNES AKBARI received the B.Sc. degree in computer software engineering from Payame Noor University (Central Branch of Birjand), Iran, in 2006, the M.Sc. degree in information technology management from Payame Noor University (Central Branch of Tehran), Tehran, Iran, in 2011, and the Ph.D. degree in applied mathematics (numerical analysis) from the University of Semnan, Iran, in 2017. He was a Postdoctoral Researcher with the Department of Computer Science and Engineering, Qatar University (QU), and worked on the project of document binarization. He is a reviewer of several international journals in the field of artificial intelligence, such as *IEEE TRANSACTIONS ON CYBERNETICS*, *Pattern Recognition*, *IEEE ACCESS*, and *Artificial Intelligence Review*. His research interests include pattern recognition, neural networks, remote sensing, and document analysis. He has published several papers on these areas.

SOMAYA AL-MAADEED (Senior Member, IEEE) received the Ph.D. degree in computer science from Nottingham, U.K., in 2004. She is the currently the Coordinator of the Computer Vision and AI Research Group. She enjoys excellent collaboration with national and international institutions and industry. She is also a principal investigator of several funded research projects generating approximately five million. She has published extensively pattern recognition and delivered workshops on teaching programming for undergraduate students. She attended workshops related to higher education strategy, assessment methods, and interactive teaching. In 2015, she was elected as the IEEE Chair for the Qatar Section.



NOOR AL-MAADEED (Member, IEEE) received the Ph.D. degree in computer engineering from Brunel University, U.K., in 2014. She is currently an Associate Professor with the Computer Science and Engineering Department, Qatar University. She participated in many regional and international conferences and published an important number of research articles in prestigious peer-reviewed journals, book chapters, and conferences proceedings. She has improved the relationship

between academia and the industry by leading many research projects, both domestically and abroad, totaling over eight million QAR in her fields of specialization, such as image processing, speech, speaker recognition, intelligent pattern recognition, video-surveillance systems, and biometrics. She is a member of the First Batch Qatar Leadership Center, the Current and Future Leaders Program, Qatar University Senate, and other committees. She is also a member of various international associations, such as IET, BA, and IAENG. She participates in activities which connect her to the community, such as working with charities and volunteering in sport events. She received the following awards, such as the Qatar Education Excellence Platinum Award for new Ph.D. holders from Highness the Emir of Qatar, in 2014 and 2015, the Premium Award from IET Biometrics, in 2017, and the Barzan Award, in 2019.



AL ANOOD NAJEEB received the bachelor's degree in computer application from the Sree Narayana Institute of Technology, India, in 2019. She is currently pursuing the M.S. degree in computer science with Qatar University, Doha, Qatar. She is also working as a Research Assistant for Dr. Somaya Al-Maadeed at Qatar University. Her research interests include image processing, computer vision, and machine learning.

AFNAN AL-ALI received the Master of Science degree in computer engineering from the University of Basra, Basra, Iraq. She is currently pursuing the Ph.D. degree with Qatar University. Her research interests include machine learning, AI, computer vision, object detection and classification, and machine learning for health care.



FOUAD KHELIFI (Member, IEEE) received the Ingenieur d'Etat degree in electrical engineering from the University of Jijel, Algeria, in 2000, the Magistère degree in electronic engineering from the University of Annaba, Algeria, in 2003, and the Ph.D. degree from the School of Computer Science, in 2007. Then, he joined Queen's University Belfast, U.K., as a Research Student, in 2004. From 2008 to 2010, he held a research position with the Digital Media and Systems Research

Institute, University of Bradford, U.K., before joining the Department of Computer and Information Sciences, Northumbria University at Newcastle, U.K., as a Lecturer (an Assistant Professor), where he is currently an Associate Professor (a Reader). He has authored and coauthored over 90 publications and has successfully supervised 12 Ph.D. students. His research interests include the fields of computer vision and machine learning, image and video watermarking, image/video authentication and perceptual hashing, data hiding, image forensics and biometrics, image and video coding, and medical image analysis.



ASHREF LAWGALY received the B.Sc. degree in computer science from Garyounis University, Benghazi, Libya, in 2001, and the M.Sc. degree (Hons.) in software engineering from Northumbria University, Newcastle, U.K., in 2011. He then joined the Northumbria University at Newcastle, as a Research Student, in 2013, and received the Ph.D. degree from the Department Computer and Information Sciences (CIS), in 2017. From 2018 to 2020, he held a research position with the Department of CIS, Northumbria University. He is currently a Senior Research Assistant with the CIS Department. His research interests include image and video processing, multimedia forensics, information security, and medical image analysis.

...