IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 14, 2020, accepted January 11, 2021, date of publication January 14, 2021, date of current version January 25, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3051921

Forensic Insights From Smartphones Through Electromagnetic Side-Channel Analysis

ASANKA P. SAYAKKARA AND NHIEN-AN LE-KHAC, (Member, IEEE)
School of Computer Science, University College Dublin, Dublin 4, D04 V1W8 Ireland
Corresponding author: Nhien-An Le-Khac (an.lekhac@ucd.ie)

This work was supported by the EU DG-HOME CERBERUS Project under Grant 822015.

ABSTRACT The increasing use of smartphones has increased their presence in legal and corporate investigations. Unlike desktop and laptop computers, forensic analysis of smartphones is a challenging task due to their limited interfaces to retrieve information of forensic value. Electromagnetic side-channel analysis (EM-SCA) has been recently proposed as an alternative window to acquire forensic insights from computers, in particularly from Internet of Things devices. Along this line, this work experimentally evaluates the potential of extracting information of forensic value from smartphones through their EM radiation. Initially, a group of smartphones representing a diverse set of system-on-chip (SoC) processors were used to acquire EM radiation traces. Later, deep learning models were trained to detect various internal software behaviours running on the SoCs. The results of this work indicates that a wide variety of insights can be extracted from smartphones through EM side-channel, increasing the potential opportunities for digital forensic investigators.

INDEX TERMS Digital forensics, smartphone forensics, electromagnetic side-channel, software behaviour detection, deep learning model.

I. INTRODUCTION

Digital forensics is the field where legal and corporate investigations are assisted with digital sources of evidence. A multitude of subdomains exists under the umbrella of digital forensics, such as file system forensics, network forensics, database forensics, and mobile device forensics [1]. In comparison to desktop and laptop computers, smartphones are more ubiquitous in day-to-day human lives. Therefore, in any legal or corporate investigation scenarios, it is highly likely to encounter smartphones as an evidence source, even when general purpose computers are not available as an evidence source [2].

Investigation on smartphones is a challenging task due to multiple reasons. When following the classical digital forensic investigation approach, it is necessary to acquire images of internal storage of the smartphones [3]. The acquisition of a forensic image from the non-volatile storage of a smartphone is impossible without rooting/jailbraking it. Such actions involve the risk of tampering the device to an unrecoverable state. Furthermore, modern smartphone operating systems use encryption to protect their internal storage [4]. Due to this reason, acquired forensic images of non-volatile storage can potentially render unusable in an investigation.

Under these circumstances, it is reasonable to perform live inspection of smartphones to identify suspicious activities as soon as a device is taken into custody. However, modern smartphones use strong user authentication mechanisms such as PIN codes, log-in patterns, and biometrics, such as fingerprint and facial recognition. Furthermore, the sheer diversity of makes and models currently in the smartphone market causes difficulty in following a unified approach to perform live analysis on smartphones in investigations [5].

Electromagnetic (EM) radiation caused by internal electronic components of computers has long been recognised to be leaking information. EM side-channel analysis (EM-SCA) is a domain that utilises a large collection of methods and algorithms to exfiltrate sensitive information from computers through their EM radiation [6]. Due to the non-invasive nature of EM-SCA, it has been proposed to be used as a forensic insight-gathering method. Various types of forensic insights have been demonstrated to be acquirable from IoT devices in the literature [7], [8]. This work explores the potential of utilising EM radiation of smartphones as a method to acquire forensic insights from them during triage examination

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

and live analysis of an investigation. This article makes the following contributions:

- Experimentally demonstrates the potential of using deep learning neural networks to extract forensic insights from smartphones through their EM radiation.
- Identifies the emerging challenges in EM side-channel forensic insight acquisition from smartphones due to the advancements of SoC processor technology.
- Proposes potential avenues for future research in order to enable EM side-channel forensic insight acquisition for smartphones in real-world investigations.

The rest of this article is organised as follows. Section II describes related literature on smartphone forensics and the use of EM-SCA in similar contexts. Section III introduces technical preliminaries that are necessary to understand the approach to acquire, preprocess and analyse EM data using specific hardware and software tools. Section IV illustrates the specific details on dealing with EM radiation from smartphones. The methods and results of acquiring forensic insights from smartphones are presented in Section V followed by a discussion on the findings in Section VI. The Section VII introduces emerging challenges and future trends in the field of EM side-channel forensics on smartphones. Finally, Section VIII concludes this article.

II. RELATED WORK

The unique identification and fingerprinting of computing devices through their side-channel radiation has been explored in the literature producing various techniques. Radio frequency distinct native attributes (RF-DNA) is a method to create a unique sequence of numbers representing a specific device based on the EM radiation it produces [9], [10]. Furthermore, the magnetic field caused by smartphones are shown to be detectable with the help of built-in magnetic sensors of another smartphone in order to fingerprint such devices [11]. Such device fingerprinting techniques can be used in forensic investigation contexts to recognise a particular smartphone make and model before proceeding to extract forensic insights.

The identification of running software activities on computing devices has been previously demonstrated on multiple work. For example, it is possible to identify cryptographic operations [8], sorting algorithms [12], and specific known software behaviours [13] on embedded systems such as Arduino and Raspberry Pi devices. Chawla *et al.* performed application inference on a SoC processor running Android operating system by utilising a combination of EM radiation and dynamic voltage frequency scaling (DVFS) information from the CPU driver [14]. In order to acquire forensic insights from smartphones in digital forensic scenarios, it is important to explore the potential of using such methods under realistic scenarios on actual smartphones.

Smartphones can be actively running audio or video steaming applications by the time they were seized. The identification of such devices during the triage examination phase can help law-enforcement to make timely interventions. Yilmaz *et al.* explored the potential of detecting smartphone camera status through their EM radiation patterns [15]. Their work acquires EM radiation coming from the system-on-chip (SoC) clock frequency of each considered smartphone type and later uses two-stage dimensionality reduction approach to select a feature vector. Finally, a classifier developed using k-Nearest Neighbors (k-NN) algorithm was used detect the smartphone model and the camera state with over 97% accuracy. An important finding of their work is the potential of detecting smartphone camera status from both near-field and far-field. The near-field EM data acquisition is performed using an H-loop probe placed over the device camera, whereas far-field EM data acquisition is performed using a planar antenna placed 5 metres away from the device.

Moving further, Yilmaz *et al.* used convolutional neural networks (CNN) to classify smartphone models and their camera status through their EM radiation [16]. In this work, four smartphone models and three camera states, i.e., front camera active, rear camera active, and camera idle were considered. Similar to the previous work, the CNN classifier achieves almost 100% accuracy in determining the device model and camera status. As the authors use a single dataset acquired at a specific time period for training and testing ML models, the generalisability of ML models to new datasets from other devices of the same tested types needs to be evaluated.

III. TECHNICAL PRELIMINARIES

This section initially introduces the theoretical and technical details behind the production and preprocessing of EM data in order to produce datasets that can be used with machine learning algorithms. Then, the theoretical and technical details of using deep learning neural network algorithms with EM datasets is presented.

A. ELECTROMAGNETIC RADIATION DATA

The acquisition of EM side-channel radiation data can be performed using different types of data acquisition tools. Among them, software-defined radio (SDR) platforms provide a great flexibility as they are both software-controllable and software-programmable [17]. SDR platforms use the complex IQ sampling with extremely fast sample rates. That means, they produces a stream of samples where each sample consists of two components, namely the in-phase (I) component and the quadrature-phase (Q) component.

When observing EM radiation from a source, such as a smartphone, using an SDR platform for a time period of T_s with a sample rate of F_s, the total number of IQ data samples produced is N_s given by the Equation 1.

$$N_s = T_s \times F_s \tag{1}$$

The conversion of a complex IQ sample EM trace, Y, into a real-valued sample EM trace, x, can be done by taking the absolute value of each complex sample as shown in the

Equation 2.

$$x_i = \sqrt{Y_{iI}^2 + Y_{iQ}^2}, \text{ where } i \in [1, N_s] \tag{2}$$

However, the EM trace, x is still in the time domain. Time domain signals are susceptible to external noise sources, which cause sudden fluctuations in the observed signal. Such effects make it difficult to recognise the useful signal patterns that represent important internal behaviours of the target smartphone. Therefore, it is desirable to handle EM data in the frequency domain. This is achieved by calculating short-time Fourier Transform (STFT) [18] as follows, which produces a matrix X:

$$STFT\{x[n]\} \equiv X(m, \omega) = \sum_{n=1}^{N_s} x[n]\omega[n-m]e^{-j\omega n} \tag{3}$$

In the Equation 3 on STFT calculation, ω represents the frequency variable, while m represents the time index of each Fourier Transform operation window. The resulting STFT matrix, X, has a frequency axis with a length of W_s. This length is equal to the size of the Fourier Transform window in the STFT operation. Meanwhile, the time axis of the STFT matrix, X, has a length of N_w, which is the number of windows produced by the STFT operation. The value of N_w can be calculated as shown in the Equation 4. There, N_s is the number of samples in the original signal, W_s is the FFT window size, and O_s is the number of overlapping samples.

$$N_w = floor\left(\frac{N_s - W_s}{W_s - O_s}\right) + 1 \tag{4}$$

$$X = \overbrace{\begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ . & . & . & . & . & . \end{bmatrix}}^{\text{Window Size } (W_s)} \Bigg\} \text{Time Steps } (N_w)$$

The X matrix ($N_w \times W_s$) resulted from an acquired EM trace serves as the basis to create datasets to train and test machine learning models.

B. DEEP LEARNING NEURAL NETWORKS

Deep learning neural networks are built by stacking multiple layers of artificial neurons on top of each other where the two outer layers take the input data and produce predictions respectively. Equation 5 represents the activation of the first layer of a neural network. X is the matrix of the input features where each row is a training instance and each column is a feature; $W^{(1)}$ is the weight matrix representing the connections from the input feature vector to each perceptron in the given layer;vector $b^{(1)}$ contains the bias weights for the connections between bias neurons and each neuron in the layer. Meanwhile, ϕ is the activation function that computes the output matrix $Z^{(1)}$ from the layer.

$$Z^{(1)} = \phi(XW^{(1)} + b^{(1)}) \tag{5}$$

Multiple metrics can be used to test the performance of a trained model using the number of true positives (TP), false positives (FP), and false negatives (FN) produced by the network on the testing dataset. The *precision* (see Equation 6) measures the accuracy of the positive predictions of the network, while the *recall* (see Equation 7) measures the ratio of making correct positive predictions against the total number of positive labels in the dataset. A better metric that combines both the precision and the recall to produce a single measure is F_1 score shown in the Equation 8.

$$precision = \frac{TP}{TP + FP} \tag{6}$$

$$recall = \frac{TP}{TP + FN} \tag{7}$$

$$F_1 = 2 \times \left(\frac{precision \times recall}{precision + recall}\right) \tag{8}$$

IV. SMARTPHONE ELECTROMAGNETIC RADIATION

For the acquisition of EM data, a HackRF One SDR device is used on this work [19]. HackRF has a tunable frequency range from 1 MHz to 6 GHz and supports sample rates up to 20 MHz. In order to capture EM radiation of smartphones at the near-field, an H-Loop near-field probe with a diameter of 25 mm is connected to the HackRF device [20]. Finally, the HackRF device is connected to a host computer via a USB cable, which runs GNU Radio library [21]. The GNU Radio library saves EM data to files as raw IQ samples where each component of a complex IQ sample is a 32 bit floating-point value. Therefore, a single IQ sample generated through this data acquisition setup is 8 bytes long.

A smartphone can be producing EM radiation from various internal hardware components in various signal frequencies. As the information related to the ongoing activities of a smartphone is associated with its SoC chip, the system clock frequency of the SoC chip can be considered as the most important EM frequency to observe [15]. In order to locate the most appropriate location to observe EM radiation from the exterior of a smartphone, the H-loop near-field probe can be moved across the surface of the device while plotting the radiation pattern as a spectrogram. The location of the strongest signal observation at the system clock frequency can be fixed as the appropriate antenna placement for EM data acquisition for a particular smartphone.

The Figure 1 illustrates the placement of the H-loop near-field probe over an iPhone 4S device for acquiring EM radiation from the device. While the smartphone is powered on and the display is active, EM radiation was observed for a brief period of time at the clock frequency of the iPhone 4S, which is 1 GHz, using a sample rate of 20 MHz. As can be seen from the spectrogram of the signal (see Figure 2), multiple peaks can be observed around 1 GHz frequency. The patterns of these peaks varies according to the user interactions made with smartphone by openning and closing different apps. Furthermore, the radiation strength of the observed signal significantly weakens when the device display goes off, which puts most of the internal software activities to a halt.

TABLE 1. The SoC chips, their system clock frequencies of different core clusters, and some examples for smartphones that employ the SoC. The star (*) symbol denotes the smartphone products that were used for the experimentation in this work as a representative of the corresponding SoC.

System-on-Chip	Architecture	CPU Frequency 1	CPU Frequency 2	Devices
Apple A5	ARMv 7-A	1 GHz (2 cores)	N/A	iPhone 4S*, iPad 2, Apple TV (3rd generation), iPod Touch (5th generation), iPad Mini (1st generation).
Qualcomm Snapdragon MSM8260A	ARM v7-A	1.5 GHz (2 cores)	N/A	Sony Xperia T*, Xiaomi Mi-2A, HTC One S, HTC Windows Phone 8X, etc.
Qualcomm Snapdragon MSM8916	ARMv 8-A	1.2 GHz (4 cores)	N/A	Samsung Galaxy Grand Prime*, Samsung Galaxy E7, Huawei G621, HTC Desire 510, etc.
Qualcomm Snapdragon SDM439	ARMv 8-A	1.95 GHz (4 cores)	1.45 GHz (4 cores)	Nokia 4.2*, Redmi 7A, Redmi 8A, Redmi 8A Dual, Samsung Galaxy A01, Vivo Y95, etc.

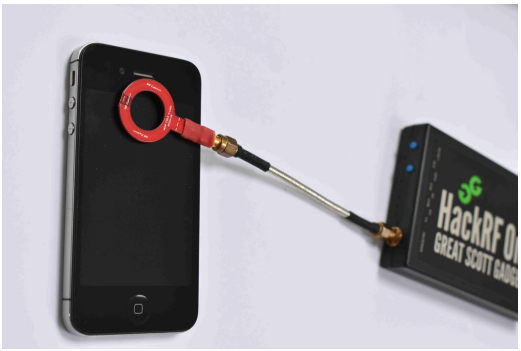

FIGURE 1. Placement of the H-loop near-field antenna of the HackRF SDR over an iPhone 4S device for EM data acquisition.

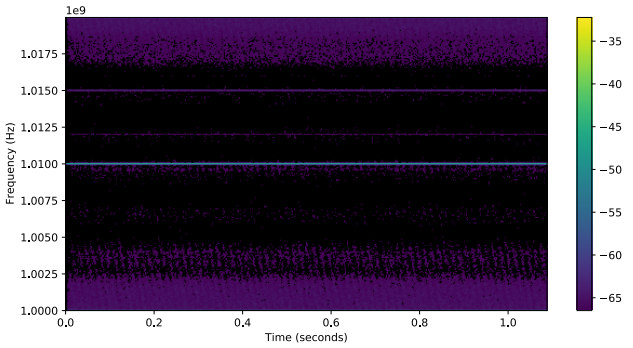

FIGURE 2. EM radiation spectrum of the iPhone 4S device around the clock frequency of its SoC processor, i.e., 1 GHz, which is observed for a time period of about 1 second.

There is a huge variety of smartphone makes and models currently in use. However, there are only a limited number of SoC chip instruction set architectures and chip manufacturers for mobile devices. Therefore, many smartphones from different manufacturers employ the same chip architectures. As a result, the methods to detect software behaviours of certain makes and models of smartphones can be potentially extended to be used on many other smartphone makes and models, as far as they use similar SoC chips. The Table 1 illustrates the type of SoC chips and their system clock frequencies from four different smartphone models. The experimental evaluations of the methods presented in this work use these four types of smartphones as EM radiation sources.

V. FORENSIC INSIGHTS FROM SMARTPHONES

This section presents the procedure to acquire and preprocess EM data from smartphones. Then, the use of preprocessed EM data with deep neural network models to extract forensic insights will be illustrated.

A. PREPARATION OF DATASET

For the purpose of this study, 4 models of smartphones are considered as target devices, which are illustrated in the Table 1. For each device, the system clock frequency of the SoC processor is considered as the target information-leaking frequency. For the SoCs that has multiple clock frequencies (CPU frequency 1 and 2 specified in Table 1), all those frequencies and their nearby harmonics were considered as potential information leaking frequencies during initial inspection of the SoCs to identify a suitable frequency.

In order to identify the location on a device where the EM radiation is the strongest, The H-loop antenna was moved around the surface of the device while visualising the clock-frequency signal in real-time as spectrogram. This process was repeated to identify best antenna position for each device. As the next step, each device was inspected to identify a set of important apps and behaviours. For the purpose of this research, the number of activities for each device was fixed to 10, which includes various apps being active on the device.

In order to acquire EM radiation data for each device, the following procedure was used. While a specific previously identified software behaviour is active on a particular device, the H-loop antenna was placed over the device and capture EM radiation data using the HackRF device for a few seconds. The EM radiation was sampled at 20 MHz for each case. This procedure resulted in 10 EM trace files for each device; 40 EM trace files in total for the 4 smartphone

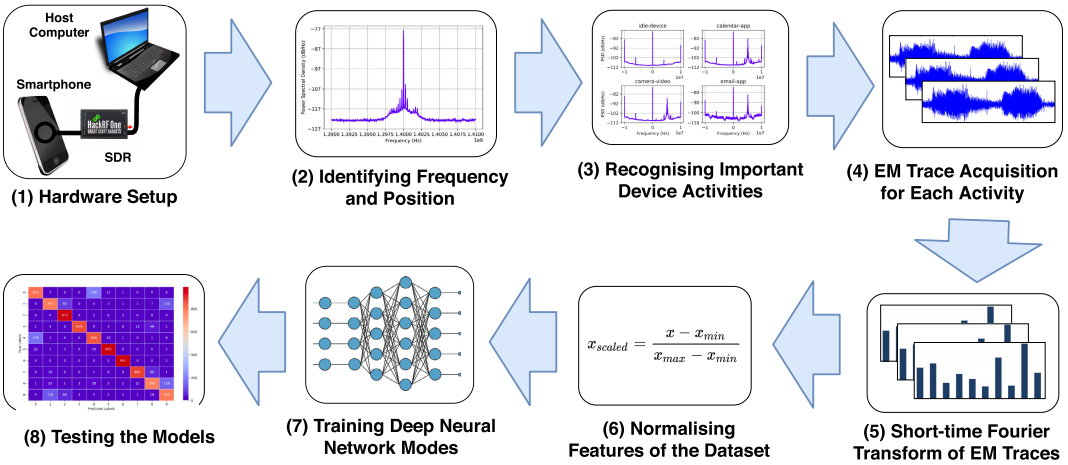

FIGURE 3. The pipeline for EM data acquisition, preprocessing, and finally training deep learning models to predict internal software activities of smartphones.

devices. Each EM trace file, which consists of time domain signal, was applied to STFT function to produce a series of frequency domain windows. For deep learning, each of such window is considered as a training instance whereas the corresponding software activity of the smartphone is considered as the label.

The resulting EM dataset of each smartphone was used to build individual deep learning models to identify software activities of the corresponding device. Some of the hyperparameters, such as input layer dimensions, of the deep learning models depend on the dimensions of the EM dataset. Therefore, the specific settings of the STFT operations, namely, the FFT window size, W_s, and the number of overlapping samples, O_s, were adjusted during the hyperparameter tuning of deep learning models as needed. Figure 3 illustrates the pipeline of data acquiring, preprocessing, and training deep learning models. Once the models are tested, they can be used in forensic investigation situations with EM data acquired from devices under investigation.

B. SYSTEM-ON-CHIP: APPLE A5

The Apple A5 is a SoC that consists of 2 cores running at 1 GHz clock frequency. It is being used on multiple devices, such as iPhone 4S, iPad 2, Apple TV (3rd generation), iPod Touch (5th generation) and iPad Mini (1st generation). For the purpose of experimentation, this work uses the iPhone 4S as a representative device. During the initial inspection of iPhone 4S device, 10 different device behaviours were identified for the deep learning-based detection through EM radiation. These device behaviours are namely, calendar-app, camera-photo, camera-video, email-app, gallary-app, home-screen, idle-device, phone-app, sms-app, and web-browser. These behaviours are labeled from 0 to 9. Figure 4 depicts the power spectral density (PSD) of the captured EM radiation for some of the software behaviours.

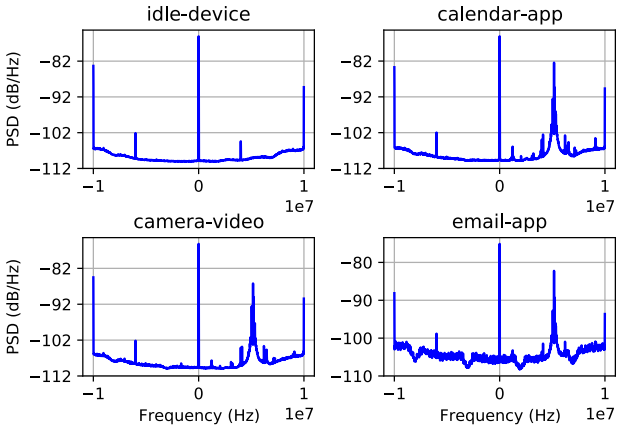

FIGURE 4. Power spectral density (PSD) of 4 different software bahavours of the iPhone 4S device.

In order to convert the time domain EM traces to the trainable dataset, the STFT window size, W_s and the overlap size, O_s need to be set. The larger the W_s value, the more frequency bandwidth is included in a single training sample. However, at the same time, it increases the dimensionality of the input samples for deep learning. Therefore, it is desirable to set the W_s to as smaller value as possible as long as it provides a sufficient classification accuracy in the deep learning models. Meanwhile, larger W_s values makes the number of available training samples smaller. To overcome this, O_s value can be increased. Considering these factors, the STFT window size, W_s, was set to 2048 along with an overlap size, O_s, of 256 by empirically testing different potential settings against the classification accuracy they produces.

When setting up the hyperparameters for a deep neural network, the input feature vector size is set to 2048 as each STFT window is of the same size. The Table 2 details the structure of the deep neural network. It consists

TABLE 2. The hypterparameters of the deep neural network setting used for iPhone 4S and Sony Xperia and Galaxy Grand Prime devices.

Layer Type	Output Shape	# of Parameters
Dense (ReLU)	1400	2868600
Dense (ReLU)	800	1120800
Dense (ReLU)	500	400500
Dense (ReLU)	200	100200
Dense (ReLU)	100	20100
Dense (Softmax)	10	1010
Total Parameters		4,511,210

TABLE 3. Performance of the deep learning model to detect 10 software behaviours of the iPhone 4S device on the testing dataset.

Class	Precision	Recall	F1-score	Support
calendar-app (0)	0.82	0.80	0.81	1011
camera-photo (1)	0.82	0.79	0.80	1010
camera-video (2)	0.86	0.98	0.92	993
email-app (3)	0.99	0.93	0.96	939
gallery-app (4)	0.79	0.82	0.81	1043
home-screen (5)	0.95	0.96	0.95	1027
idle-device (6)	0.99	0.99	0.99	999
phone-app (7)	0.97	0.91	0.94	974
sms-app (8)	0.85	0.78	0.82	1003
web-browser (9)	0.74	0.79	0.77	1001
Macro Avg	0.88	0.88	0.88	10000
Weighted Avg	0.88	0.87	0.87	10000
Accuracy			0.87	10000

of 5 hidden dense layers with decreasing number of dimensions. The hidden layers are configured to use Rectified Linear Unit (ReLU) as the activation function. The output layer consists of 10 nodes as there are 10 classes to predict and uses Softmax as the activation function. The ReLU activation function produces a linear output for non-zero inputs making it suitable for internal layers, whereas Softmax activation function helps to produce a probability distribution across the output nodes making it suitable for predicting classes with probabilities.

In order to create the dataset for training the deep neural network, 10,000 training instances, i.e., STFT windows, were taken from each class, which results in a total of 100,000 training instances for the 10 classes. In the beginning, the training and testing samples were separated by 9:1 ratio. Then, the network was trained for 50 epochs where a random 10% of the training dataset is used for validation at the end of each epoch. The network uses stochastic gradient descent (SGD) as the cost/optimisation function with a learning rate of 0.001 and sparse categorical crossentropy as the loss function.

Table 3 illustrates the performance of the trained deep learning model on the testing dataset. Macro average of the precision, recall and F1 metrics represents the average value acquired for each class, whereas weighted average represents the average value of the metrics for each class adjusted according to the number of testing samples available for each class, i.e., support values in the table. As it is evident from the table, the trained model achieves an accuracy of 87% on testing data. Figure 5 depicts the confusion matrix of the deep learning model for testing data.

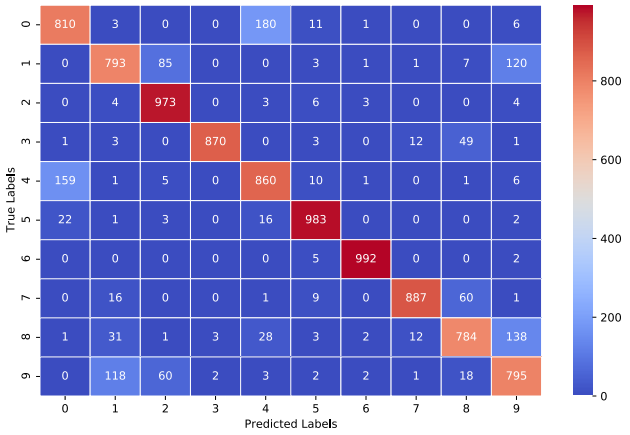

FIGURE 5. Confusion matrix of the deep learning model to detect 10 software behaviours of the iPhone 4S device on the testing dataset.

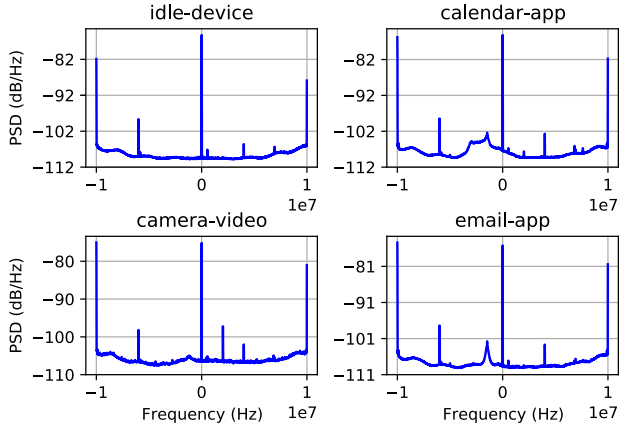

FIGURE 6. Power spectral density (PSD) of 4 different software behaviours of the Sony Xperia device.

C. SYSTEM-ON-CHIP: QUALCOMM SNAPDRAGON MSM8260A

The Qualcomm Snapdragon MSM8260A is a SoC with 2 cores running at 1.5 GHz clock frequency. The SoC it being used on multiple smartphones in the market, such as Sony Xperia T, Xiaomi Mi-2A, HTC One S, and HTC Windows Phone 8X. For the experimentation of this work, the Sony Xperia T device was used as a representative of the SoC where the same set of device behaviour activities of the Apple iPhone 4S were selected. After identifying the leakage frequency and the best antenna position for the Sony Xperia device, 10 EM trace files were acquired representing the 10 device behavioural states. The Figure 6 illustrates the PSD of some of the software bahaviour classes of the Sony Xperia device. The EM traces were preprocessed and used to create a dataset according the same procedure described previously in Subsection V-B for iPhone 4S device. The

TABLE 4. Performance of the deep learning model to detect 10 software behaviours of the Sony Xperia device on the testing dataset.

Class	Precision	Recall	F1-score	Support
calendar-app (0)	0.97	0.96	0.96	1011
camera-photo (1)	0.94	0.97	0.95	1010
camera-video (2)	0.84	0.88	0.86	993
email-app (3)	0.79	0.90	0.84	939
gallery-app (4)	0.84	0.69	0.76	1043
home-screen (5)	0.77	0.73	0.75	1027
idle-device (6)	0.94	0.97	0.96	999
phone-app (7)	0.94	0.95	0.95	974
sms-app (8)	0.79	0.81	0.80	1003
web-browser (9)	0.98	0.97	0.98	1001
Macro Avg	0.88	0.88	0.88	10000
Weighted Avg	0.88	0.88	0.88	10000
Accuracy			0.88	10000

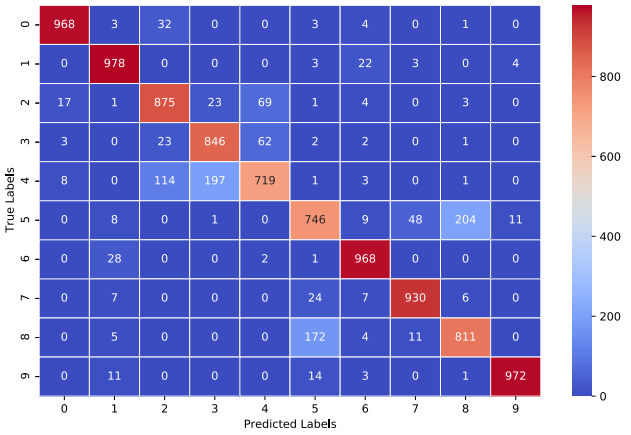

FIGURE 7. Confusion matrix of the deep learning model to detect 10 software behaviours of the Sony Xperia device on the testing dataset.

Table 2 illustrates the hyperparameter configurations for the deep neural network used to train and test smartphone device behaviour.

The Table 4 illustrates the classification results obtained from the trained deep learning model for Sony Xperia device. It is evident that the internal behaviours of the device can be distinguished with an accuracy of 88%, which is almost similar to the iPhone 4S device analysed by using a similarly configured deep learning model. Figure 7 illustrates the confusion matrix for the deep learning model of Sony Xperia device performing on training dataset.

D. SYSTEM-ON-CHIP: QUALCOMM SNAPDRAGON MSM8916

The Qualcomm Snapdragon MSM8916 is a quad-core processor with a clock frequency of 1.2 GHz. The SoC is being used on various mobile devices, such as Samsung Galaxy Grand Prime, Samsung Galaxy E7, Huawei G621, and HTC Desire 510. Among them, Samsung Galaxy Grand Prime device was selected as a representative of the SoC for the experimentation of this work. Following the same procedure for acquiring and processing EM data, a deep neural network was trained to detect software behaviours of a Galaxy Grand

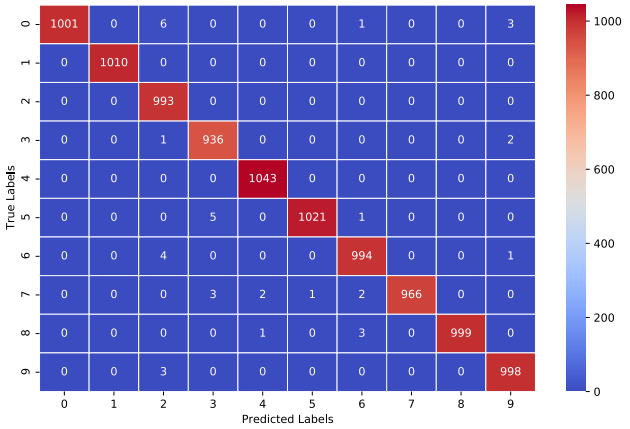

FIGURE 8. Confusion matrix of the deep learning model to detect 10 software behaviours of the Galaxy Grand Prime device on the testing dataset.

Prime device. The hyperparameters of the deep neural network was set to the same values of the previous cases (see Table 2). The model achieved an accuracy of 99% on the testing dataset for 10 different software behaviours with just 10 epochs of training. The confusion matrix is shown in the Figure 8.

E. SYSTEM-ON-CHIP: QUALCOMM SNAPDRAGON SDM439

The Qualcomm Snapdragon SDM439 is an octa-core SoC that are organised as two equally-sized clusters. The 4 cores of the first cluster runs at 1.95 GHz clock frequency, while the 4 cores of the second cluster runs at 1.45 GHz clock frequency. The SoC is currently being used on multiple mobile devices, such as Nokia 4.2, Redmi 7A, Redmi 8A, Redmi 8A Dual, Samsung Galaxy A01, and Vivo Y95. For the experimentation of this work, the Nokia 4.2 device was selected as a representative of the SoC.

The observation of EM radiation from Nokia 4.2 device was a challenging task due to the two CPU clusters with two different clock frequencies. The potential frequency range to scan through during initial inspection is wider compared to SoCs that has a single clock frequency. Even after a specific frequency was identified, there is no guarantee whether it is the only information-leaking EM radiation coming from the SoC or whether the identified frequency is sufficient to infer internal software behaviours of the device.

During the initial inspection of the device, a signal was identified at 1.53 GHz that indicated a changing pattern against the variation of software activities that were running on the device. Therefore, this frequency was used to acquire EM radiation from the device. Similar to previous scenarios, EM radiation traces were acquired while the Nokia 4.2 device was running 10 different software behaviours. However, when these EM traces were visualised to see their distinctiveness, it was noted that most EM patterns of software behaviours looked similar in PSD plots. Only a few

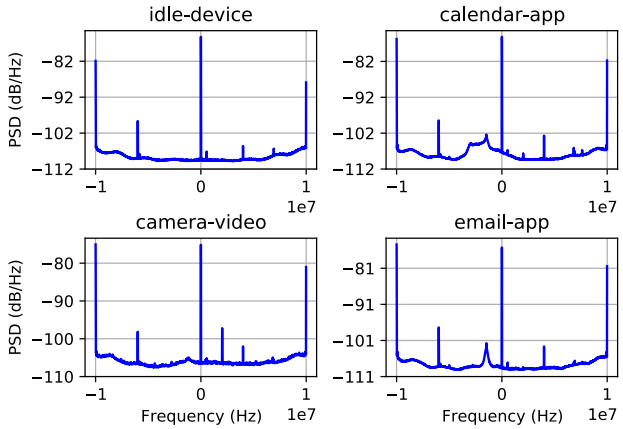

FIGURE 9. Power spectral density (PSD) of 4 different software behaviours of the Nokia 4.2 device.

TABLE 5. The hypterparameters of the deep neural network setting used for Nokia 4.2 device.

Layer Type	Output Shape	# of Parameters
Dense (ReLU)	1400	2868600
Dense (ReLU)	800	1120800
Dense (ReLU)	500	400500
Dense (ReLU)	200	100200
Dense (ReLU)	100	20100
Dense (ReLU)	50	5050
Dense (Softmax)	4	204
Total Parameters		4,515,454

activities (see Figure 9) demonstrated a noticeable pattern difference from each other. Therefore, 4 software activities, namely *idle device*, *calendar app*, *camera app*, and *email app*, were used for training and testing a deep learning model to distinguish between them.

The Table 5 illustrates the hyperparameter configuration for the deep neural network used for this particular classification. The last layer of the network was set to contain 4 nodes as there are two classes to classify and it uses a Softmax activation function. For training the network, 20,000 training instances from each class were used, totalling to 80,000 instances. Finally, with a 0.005 learning rate and 30 epochs, the network was trained. A 10% of the dataset (8000 instances) was separated at the beginning for the testing of the trained model, which produced the results illustrated in Table 6. As can be seen, the classification accuracy reached 82% on the testing data.

Further attempts were made to build deep models with 10 classes in order to distinguish between the 10 software activities of the original device. However, as noticed initially by visualisations of the data, the models could not achieve a significant classification accuracy for various neural network hyperparameter configurations considered for the 10 class classification problem. In all configurations, the networks only achieved an accuracy closer to 51% indicating that the leakage patterns are not sufficiently distinguishable.

TABLE 6. Performance of the deep learning model to detect 4 software behaviours of the Nokia 4.2 device on the testing dataset.

Class	Precision	Recall	F1-score	Support
camera-video (0)	0.99	0.99	0.99	1926
calendar-app (1)	0.72	0.78	0.75	1955
idle-device (2)	0.83	0.99	0.90	2096
web-browser (9)	0.75	0.53	0.62	2023
Macro Avg	0.82	0.82	0.82	8000
Weighted Avg	0.82	0.82	0.81	8000
Accuracy			0.82	8000

VI. DISCUSSION

For the purpose of acquiring digital evidence, various digital forensic techniques and tools have been developed. In many situations, these techniques can be extended to investigate on smartphones as well. However, the nature of smartphones consists of aspects that are not available on typical computing systems, such as desktop and laptop computers. The heterogeneity of smartphones in use from various vendors with a diverse set of hardware and software configurations make the development of forensic tools a challenging task [5]. In such dynamic environments, the application of artificial intelligence (AI) facilitates the resilience to adapt to rapidly changing requirements.

The deep learning models trained and tested in this work indicates that specific predefined software behaviours of smartphones can be extracted through their EM radiation without any invasive actions to the devices. Depending on changing requirements, new deep learning models can be trained targeted at new devices and their software behaviours. The EM radiation acquisition from a smartphones is directly focused on the EM radiation of their on-board SoC. Therefore, the deep learning models are actually trained to detect the behaviours of specific SoCs instead of smartphones. As a result, a model trained to detect software behaviours of a specific SoC can potentially generalise across many smartphone makes and models that uses the same chip. For example, although a deep learning model was trained for Qualcomm Snapdragon SDM439 SoC using Nokia 4.2 smartphone in this work, the same deep learning model can potential be used to detect software behaviours on Samsung Galaxy A01 smartphone which is still in the market. Further studies are necessary to identify the impact of device-specific factors, such as a smartphone's physical design, the orientation and location of the SoC on the device's circuit board, to the generalisability of a deep learning model across devices.

The techniques and tools developed to acquire forensic evidence from computing systems have to be *forensically sound*. According to McKemmish [22], the forensic soundness of a digital forensic investigation procedure is ensured by 4 evaluation criteria. Firstly, the interpretation of the the electronic evidence should not affected by the investigation process in question. Secondly, the potential errors and doubts that are involved with the forensic method should have been reasonably identified and satisfactorily explained. Thirdly, the analysis process should be independently verifiable by

TABLE 7. Two recent SoC chips Apple A10 and A14 that uses ARM's big.LITTLE technology to distribute workload between multiple CPU clusters with different capabilities.

System-on-Chip	Architecture	CPU Frequency 1	CPU Frequency 2	Devices
Apple A10	ARM v8-A	2.34 GHz (2 cores)	1.1 GHz (2 cores)	iPhone 7 and 7 Plus, iPad 6th and 7th generations, iPod touch (7th generation).
Apple A14	ARMv 8-A	3.1 GHz (2 cores)	1.8 GHz (4 cores)	iPad Air (4th Generation), iPhone 12 and iPhone 12 Mini, iPhone 12 Pro and iPhone 12 Pro Max.

multiple parties and should produce the same result. Finally, the analysis process should be carried out by a sufficiently experienced and skilled individual. The procedure of gathering insights about the internal behaviours of smartphones can satisfy the aforementioned first, third and fourth criteria. However, the use of deep learning models to make predictions does not satisfy the second criterion as the predictions are associated with a probability and does not ensure a 100% accuracy. This is why the information gathered through EM side-channel analysis is called forensic insights instead of forensic evidence. Consequently, these insights should be only used as hints for an investigator to proceed with an investigation and discover court-admissible forensic evidence through other means.

While EM data can be acquired through various other hardware devices such as signal analysers and oscilloscopes, SDR hardware provides a great flexibility in adjusting the EM frequency, the sample rate, the level of signal amplification, and many more for users who are not experts in radio frequency (RF) engineering. Throughout the experiments of this work, the sample rate of the SDR hardware was set to 20MHz, which is the maximum rate supported by the HackRF SDR device. However, the level of precision in sampling can be increased further by using SDR devices with much larger sample rates if necessary. Furthermore, external signal amplifiers can be used to boost weak EM radiation coming from devices, which are currently not possible to be inspected. In digital forensic investigation contexts, a smartphone can be inspected in close proximity and therefore the use of near-field H-loop antennas satisfies the requirement. However, in situations where it is not possible to get to the close proximity of a device, the use of far-field antennas can be considered as demonstrated by Yilmaz *et al.* [16].

VII. CURRENT CHALLENGES AND FUTURE TRENDS

When performing EM-SCA to a computing system to acquire forensic insights, one of the important starting steps is to identify the information-leaking frequency channels of the device. Currently, the system clock frequency of the CPU is considered as the key focus. While this is straightforward for many existing SoC processors used on smartphones, multiple trends in the domain are increasingly threatening the current procedure of acquiring EM radiation data.

Modern SoC processors used on smartphones consists of multiple cores that can run multiple threads simultaneously. This means that multiple forensically-important software activities can be running at the same time on the device, such as encrypting data stored in non-volatile storage and performing communication through wireless network. With the increasing number of cores, the number of simultaneous activities increases further. As long as the CPU cores are running at the same clock frequency, EM radiation can be observed and EM trace data can be collected. However, the classification of these data using machine learning techniques into a predefined set of classes becomes a significant challenge. This requires further research into multi-label classification methods to detect multiple activities running simultaneously [23].

Another recent trend in mobile device SoC processors is the use of various strategies to increase the performance of processing while reducing the energy consumption. One approach for that is the design of SoCs with multiple clusters of CPU cores that has different processing capabilities; each cluster of CPU cores runs at a specific clock frequency and provides a unique processing power. Mobile device SoCs such as Apple A10 and A14 (see Table 7) that are based on ARM instruction set architectures uses ARM's big.LITTLE technology that makes use of this strategy [24]. For example, Apple A10 SoC processor runs activities that are less processing intensive, such as an email app, on the 2 cores that runs at 1.1 GHz that consumes less energy from the battery. When it is required to run computationally heavy activities, such as playing a game app or a high definition video, the SoC deploys the 2 cores that runs at 2.34 GHz consuming more energy but with improved performance. Due to the need of multi-tasking, the SoC can dynamically switches between the two clusters of CPU cores, making it difficult to capture EM radiation in either of the 1.1 GHz and 2.34 GHz frequencies.

Similarly, another potential problem that adds up to this situation is the use of dynamic voltage frequency scaling (DVFS) techniques in modern processors [14]. It allows a SoC processor to dynamically adjust its clock frequencies according to workload. These techniques demands the observation of EM radiation simultaneously at multiple frequencies and also requires to dynamically change the signal observation frequency on-demand without prior information.

A potential approach to observe EM radiation from SoC processors that has such complexities is to disable others and isolate a specific CPU core by running a selected thread on it. Furthermore, DVFS can be disabled to fix the CPU

clock frequency to a predefined values. By doing so, the EM radiation coming from a specific CPU core on a specific frequency can be extracted and be used to profile it. However, in order to extract forensic insights from SoC processors in realistic scenarios, it is necessary to develop new EM data processing and deep learning methods to combine and generalise the profiles of different cores or core clusters of a SoC.

The mitigation of side-channel information leakage from computing systems is an active research area with various potential methods already published and in use [25], [26]. These methods include both software and hardware mitigation strategies, such as randomisation of software activities, physical shielding of the device's internal components, use of dual-line logic, etc. In this work, it was observed that an increasing difficulty in capturing EM radiation when moving from SoCs of ARM version 7 to version 8 instruction set architectures. The ARM version 7 consists of 32-bit instructions while ARM version 8 includes both 32 and 64 bit instructions. It is possible that the latest versions of ARM CPU designs include side-channel countermeasures that prevents the observation of EM radiation with existing methods. This situation poses a challenge to the EM side-channel analysis on mobile devices in the future and demands the development of novel methods that are resilient to side-channel mitigation techniques.

In addition to the EM side-channel analysis, there exists various other non-invasive side-channels for exfiltrating information from computing systems, such as acoustic and power side-channels [27], [28]. Certain information that are not leaked through the EM side-channel can be available on such other side-channels. Therefore, it would be interesting to explore the potential of combining multiple side-channel analysis approaches together to increase the attack surface on smartphones. In addition to the typical EM side-channel radiation that directly emits from the processor chip, a new kind of attack called *screaming channels* has been introduced where the EM radiation of CPU operations get modulated into the legitimate radio transmissions from the SoC processor, such as WiFi, Bluetooth, and cellular network transmissions [29], [30]. Future research should also explore the potential of using such extended EM side-channel attacks to extract forensic insights from smartphones from a distance.

VIII. CONCLUSION

The work presented in this article introduced EM side-channel analysis as a window to extract forensic insights from smartphones. Using empirical evaluations, it was shown that deep learning models can be trained to detect specific predefined software activities running on smartphones through the EM radiation they produce with near-field. The deep learning models achieved classification accuracies varying from 82% to 99% on different smartphones. Furthermore, this work identified the emerging challenges and opportunities to the application of EM-SCA for forensic insight acquisition from smartphones that need to be the focus of future research.

REFERENCES

[1] E. Casey, *Digital Evidence and Computer Crime: Forensic Science, Computers and the Internet*. New York, NY, USA: Academic, 2011.
[2] A. Mylonas, V. Meletiadis, B. Tsoumas, L. Mitrou, and D. Gritzalis, "Smartphone forensics: A proactive investigation scheme for evidence acquisition," in *Proc. IFIP Int. Inf. Secur. Conf.* Berlin, Germany: Springer, 2012, pp. 249–260, doi: 10.1007/978-3-642-30436-1_21.
[3] K. Barmpatsalou, T. Cruz, E. Monteiro, and P. Simoes, "Current and future trends in mobile device forensics: A survey," *ACM Comput. Surv.*, vol. 51, no. 3, pp. 1–31, Jul. 2018, doi: 10.1145/3177847.
[4] E. Casey and G. J. Stellatos, "The impact of full disk encryption on digital forensics," *ACM SIGOPS Operating Syst. Rev.*, vol. 42, no. 3, pp. 93–98, Apr. 2008, doi: 10.1145/1368506.1368519.
[5] M. Chernyshev, S. Zeadally, Z. Baig, and A. Woodward, "Mobile forensics: Advances, challenges, and research opportunities," *IEEE Secur. Privacy*, vol. 15, no. 6, pp. 42–51, Nov. 2017, doi: 10.1109/MSP.2017.4251107.
[6] E. Peeters, F.-X. Standaert, and J.-J. Quisquater, "Power and electromagnetic analysis: Improved model, consequences and comparisons," *Integration*, vol. 40, no. 1, pp. 52–60, Jan. 2007, doi: 10.1016/j.vlsi.2005.12.013.
[7] A. Sayakkara, N.-A. Le-Khac, and M. Scanlon, "A survey of electromagnetic side-channel attacks and discussion on their case-progressing potential for digital forensics," *Digit. Invest.*, vol. 29, pp. 43–54, Jun. 2019, doi: 10.1016/j.diin.2019.03.002.
[8] A. Sayakkara, N.-A. Le-Khac, and M. Scanlon, "Leveraging electromagnetic side-channel analysis for the investigation of IoT devices," *Digit. Invest.*, vol. 29, pp. S94–S103, Jul. 2019, doi: 10.1016/j.diin.2019.04.012.
[9] C. K. Dubendorfer, "Using RF-DNA fingerprints to discriminate ZigBee devices in an operational environment," M.S. thesis, Air Force Inst. Technol., Wright-Patterson Air Force Base, OH, USA, 2013.
[10] R. D. Deppensmith and S. J. Stone, "Optimized fingerprint generation using unintentional emission radio-frequency distinct native attributes (RF-DNA)," in *Proc. IEEE Nat. Aerosp. Electron. Conf.*, Jun. 2014, pp. 327–330, doi: 10.1109/naecon.2014.7045829.
[11] B. Perez, M. Musolesi, and G. Stringhini, "Fatal attraction: Identifying mobile devices through electromagnetic emissions," in *Proc. 12th Conf. Secur. Privacy Wireless Mobile Netw.*, May 2019, pp. 163–173, doi: 10.1145/3317549.3319726.
[12] A. Sayakkara, L. Miralles-Pechuán, N.-A. Le-Khac, and M. Scanlon, "Cutting through the emissions: Feature selection from electromagnetic side-channel data for activity detection," *Forensic Sci. Int., Digit. Invest.*, vol. 32, Apr. 2020, Art. no. 300927, doi: 10.1016/j.fsidi.2020.300927.
[13] A. Sayakkara, N.-A. Le-Khac, and M. Scanlon, "Facilitating electromagnetic side-channel analysis for IoT investigation: Evaluating the EMvidence framework," *Forensic Sci. Int., Digit. Invest.*, vol. 33, Jul. 2020, Art. no. 301003, doi: 10.1016/j.fsidi.2020.301003.
[14] N. Chawla, A. Singh, M. Kar, and S. Mukhopadhyay, "Application inference using machine learning based side channel analysis," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2019, pp. 1–8, doi: 10.1109/ijcnn.2019.8852124.
[15] B. B. Yilmaz, E. M. Ugurlu, M. Prvulovic, and A. Zajic, "Detecting cellphone camera status at distance by exploiting electromagnetic emanations," in *Proc. MILCOM-IEEE Mil. Commun. Conf. (MILCOM)*, Nov. 2019, pp. 1–6, doi: 10.1109/milcom47813.2019.9021060.
[16] B. B. Yilmaz, E. Mert Ugurlu, A. Zajic, and M. Prvulovic, "Cell-phone classification: A convolutional neural network approach exploiting electromagnetic emanations," in *Proc. ICASSP-IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2020, pp. 2862–2866, doi: 10.1109/icassp40776.2020.9054006.
[17] A. M. Wyglinski, R. Getz, T. Collins, and D. Pu, *Software-Defined Radio for Engineers*. Norwood, MA, USA: Artech House, 2018.
[18] J. Allen, "Short term spectral analysis, synthesis, and modification by discrete Fourier transform," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-25, no. 3, pp. 235–238, Jun. 1977, doi: 10.1109/tassp.1977.1162950.
[19] M. Ossmann. *HackRF*. Accessed: Aug. 29, 2020. [Online]. Available: https://greatscottgadgets.com/hackrf/

[20] RF Explorer, Spain. (2017). *RF Explorer Near Field Antenna Kit Datasheet*. [Online]. Available: http://j3.rf-explorer.com
[21] E. Blossom, "GNU Radio: Tools for exploring the radio frequency spectrum," *Linux J.*, vol. 2004, no. 122, p. 4, 2004.
[22] R. McKemmish, "When is digital evidence forensically sound?" in *Proc. IFIP Int. Conf. Digit. Forensics*. Boston, MA, USA: Springer, 2008, pp. 3–15, doi: 10.1007/978-0-387-84927-0_1.
[23] F. Herrera, F. Charte, A. J. Rivera, and M. J. D. Jesus, "Multilabel classification," in *Multilabel Classification*. Cham, Switzerland: Springer, 2016, pp. 17–31.
[24] *ARM Big. LITTLE Technology*. Accessed: Sep. 12, 2020. [Online]. Available: https://www.arm.com/why-arm/technologies/big-little
[25] J.-J. Quisquater and D. Samyde, "Electromagnetic analysis (EMA): Measures and counter-measures for smart cards," in *Smart Card Programming and Security*. Berlin, Germany: Springer, 2001, pp. 200–210, doi: 10.1007/3-540-45418-7_17.
[26] A. Zankl, H. Seuschek, G. Irazoqui, and B. Gulmezoglu, "Side-channel attacks in the Internet of Things: Threats and challenges," in *Solutions for Cyber-Physical Systems Ubiquity*. Hershey, PA, USA: IGI Global, 2018, pp. 325–357.
[27] D. Genkin, A. Shamir, and E. Tromer, "RSA key extraction via low-bandwidth acoustic cryptanalysis," in *Proc. Int. Cryptol. Conf.* Berlin, Germany: Springer, 2014, pp. 444–461, doi: 10.1007/978-3-662-44371-2_25
[28] P. Kocher, J. Jaffe, B. Jun, and P. Rohatgi, "Introduction to differential power analysis," *J. Cryptograph. Eng.*, vol. 1, no. 1, pp. 5–27, Apr. 2011, doi: 10.1007/s13389-011-0006-y.
[29] G. Camurati, S. Poeplau, M. Muench, T. Hayes, and A. Francillon, "Screaming channels: When electromagnetic side channels meet radio transceivers," in *Proc. 25th ACM Conf. Comput. Commun. Secur. (CCS)*, Oct. 2018, pp. 163–177, doi: 10.1145/3243734.3243802.
[30] G. Camurati, A. Francillon, and F.-X. Standaert, "Understanding screaming channels: From a detailed analysis to improved attacks," *IACR Trans. Cryptograph. Hardw. Embedded Syst.*, vol. 2020, no. 3, pp. 358–401, 2020, doi: 10.13154/tches.v2020.i3.358-401.

ASANKA P. SAYAKKARA received the B.Sc. degree from the School of Computing, University of Colombo, Sri Lanka, in 2013, and the Ph.D. degree in computer science from the University College Dublin (UCD), Ireland, in 2020. He is currently a Research Assistant with the School of Computer Science, UCD. His research interests include the Internet of Things, digital forensics, and security.

NHIEN-AN LE-KHAC (Member, IEEE) received the Ph.D. degree in computer science with the Institut National Polytechnique Grenoble, France, in 2006. He is currently a Lecturer with the School of Computer Science, University College Dublin (UCD), Ireland. He is also the Programme Director of the M.Sc. Program in forensic computing and cybercrime investigation. He is also the Co-Founder of the UCD-GNECB Postgraduate Certificate in fraud and e-crime investigation. Since 2013, he has been collaborating on many research projects as a principal/co-PI/funded investigator. He has authored or coauthored more than 150 scientific articles in peer-reviewed journals and conferences in related research fields. His research interest spans the area of cybersecurity and digital forensics, machine learning for cyber security, fraud and criminal detection, cloud security and privacy, high-performance computing, and knowledge engineering.