Lightweight Cryptographic Models for IoT Devices: A Deep Learning Approach to Power Side-Channel Attack Prevention

Luheb K. Qurban

College of Education for Pure Science, Ibn Al-Haitham, University of Baghdad, Baghdad, 00964, Iraq.

DOI: 10.37648/ijps.v19i01.005

[1]*Received: 19/01/2025; Accepted: 05/03/2025; Published: 08/03/2025*

ABSTRACT

The implementation of lightweight cryptography is often found in unrolled architecture, which offers the advantages of low latency and high real-time performance but also runs the risk of Side-Channel Attack (SCA). These days, the Internet of Things has led to a variety of applications that need lightweight cryptographic primitives, including block cyphers, for safe and effective computation with little resources. The expense of developing machine learning (ML) models makes them potentially trade secrets. They must thus be shielded against harmful types of reverse engineering (such as in IP piracy). As machine learning continues to move to edge devices, partly for performance reasons and partly for privacy reasons, the models are now vulnerable to what are known as physical side-channel assaults. Earlier studies have shown that power-based side-channel assaults may recover such control flow in highly restricted contexts, but they depended on significant changes in computational stages or data dependencies to differentiate between states in a state machine. Using Field Programming Gate Arrays (FPGAs), we investigated possible security vulnerabilities involving side-channel assaults (SCAs) based on power analysis. We have significantly improved our study report in three ways. The power analysis or power profile of FPGA, which depends on the leakage of voltage fluctuations during certain encryption activities, was covered first. A physical source, such as an oscilloscope, or a remote source, such as delay line sensors, are used to detect the fluctuations in voltage of the cryptography module. Second, we spoke about possible power analysis-based SCAs that extracted the secret key using these voltage fluctuation readings. Third, we have created a framework for successful assaults and secret key predictions that is based on machine learning (ML) and deep learning (DL) algorithms. First off, using only 570 attack power traces, our proprietary convolutional neural networks (CNN) model successfully executed an attack and exposed all 16 bytes of the secret key. Second, the same architecture has been used to effectively attack the multi-layer perceptron (MLP) model using only 3200 traces. In terms of training time, prediction time, attack time, and the amount of power traces needed for a successful attack, we have improved overall.

Keywords: *Side-Channel Attack (SCA); Multi-Layer Perceptron (MLP) Model; Field Programmable Gate Arrays (FPGAs); Machine Learning (ML); Side-Channel Attacks (SCAs); Lightweight Cryptography; Prediction Time; Secret Key; Power Traces.*

INTRODUCTION

Attackers may access physical equipment, hence cryptographic techniques in hardware guard against both physical and mathematical cryptanalysis [1]. This provides defence against threats in both areas. Physical assaults known as "side-channel attacks" exploit data that has been leaked from cryptographic hardware to compromise a security system. Examples include data on temperature, power consumption, noise, runtime, and electromagnetic radiation [2, 3]. Side-channel assaults against tangible goods, such as transportation cards and phones, are become more frequent and potent [3–5].

Side-channel attacks may be classified as either profiled or non-profiled depending on the attacker's working environment. A profiled assault, like the Template assault or the Stochastic Attack, is an example of SCA. A profiling device that is architecturally similar to the gadget that will be the target of the attack is used, along with a fixed secret key. Prior to using the acquired information to study the target device, attackers would first profile the leakage on the

[1] *How to cite the article:* Qurban L.K. (March 2025); Lightweight Cryptographic Models for IoT Devices: A Deep Learning Approach to Power Side-Channel Attack Prevention; *International Journal of Professional Studies*; Jan-Jun 2025, Vol 19, 49-61; DOI: http://doi.org/10.37648/ijps.v19i01.005

target device using profiling tools [3, 5]. Conversely, a non-profiled attack is a kind of side-channel attack that occurs in an environment devoid of profiling tools. Attackers would combine the measures they have taken from the target device with statistical techniques to assess secret keys [5, 6].

A method for protecting data flows on networks is cryptography, especially lightweight cryptography (LWC), which provides data security via digital signatures, hashing techniques, encryption, and quantum cryptography. Any amount of data may be transformed into fixed-length alphanumeric sequences using hashing methods. By transforming the data into a hash value to identify changes at the recipient's end, it preserves the data integrity. The digital signature guarantees that data is not tampered with and confirms the sender's identity. Quantum cryptography develops secure communication methods by using the concepts of quantum mechanics. They are essential for protecting data against theft, alteration, and assault [6, 7]. It protects the data's integrity, confidentiality, and authenticity against unauthorised users.

Data confidentiality is preserved via symmetric or asymmetric encryption, which hides data from unwanted access. Asymmetric encryption makes use of two separate keys: [4] a private key for decryption that is kept secret between the sender and the receiver, and a public key for encryption that is publicly published. Asymmetric encryption's dependence on big key sizes is a significant disadvantage that might result in higher processing time and complexity [11]. Although symmetric encryption is effective and suitable for encrypting huge amounts of data, it presents a security concern since the secrecy of the key must be exchanged between parties.

With low power consumption, low memory needs, and a tiny footprint that is appropriate for IoT, the LWC idea has shown promise. Lightweight block cyphers are optimised versions of classic crypto systems like Grostl-inspired fast and tiny (GIFT), PRESENT, RECTANGLE, and [19] that are derived from AES. By comparison, the LEA was designed by HIGHT. Generally speaking, they are vulnerable to attacks and have trouble becoming resilient in Internet of Things applications. Data security issues plague the Data Encryption Standard (DES), which was the industry standard for 20 years until the end of the 20th century. The National Institute of Standards and Technology (NIST) developed the Rijndael AES recommended as the standard following DES [4, 9].

Due to distinct security keys, the traditional AES encryption technique has remained resistant to cryptographic assaults even in the quantum age. High memory, traffic from the network, wide regions, and high-power consumption, however, provide difficulties for hardware implementation [9, 11]. Modern IoT projects including Lora WAN, SiGFoX, Chaos, and Zwave, as well as other IoT standards like IEEE802.11n, IEEE802.15.4, [8, 12], and ZigBee, all use AES by default. For applications with limited resources that use 8-, 16-, and 32-bit data pathways, AES is best suited for a certain set of architectural implementations [6, 9].

The internet of things (IoT) and cloud-based offerings are becoming more popular these days because of their quick computation and energy economy. However, this has brought attention to the data security and integrity problem, which is growing daily as well. Many vulnerabilities exist, including reverse engineering, software and hardware trojan horses, SCAs, [18, 19], overbuilding, intellectual property protection, etc.

In 1996, side-channel analysis was used as a means of recovering the key using time series analysis. Research on side-channel attacks based on machine learning was spurred by the later widespread usage of algorithms like the support vector machine (SVM) and random forest (RF) in the early machine learning algorithms. Dependent on whether the attacker and the attack victim are utilising the same experimental apparatus, machine learning-based side-channel assaults may be classified as either supervised or unsupervised.

Among them, SCAs are significant and pose a major risk to the use of crypto algorithms in FPGAs and other system-on-a-chip (SoC) devices [1, 9, 5]. By leveraging the physical characteristics of side channels, such as monitoring power consumption, electromagnetic radiation, or time execution while carrying out any cryptographic operation on the device, this leaks the secret bits of information from the victim cryptography device. A general graphical depiction of SCAs on a target cryptographic device is shown in Fig. 1 [7, 18].

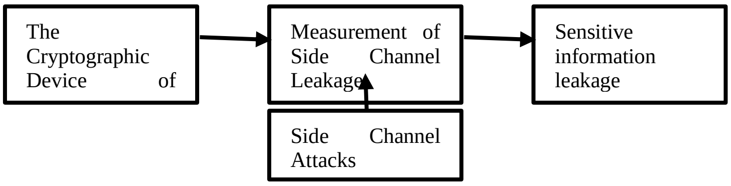
The Cryptographic Device of
Measurement of Side Channel Leakage
Sensitive information leakage
Side Channel Attacks

Fig. 1 SCAs in a general visualisation. [22]

Cyberattacks have increased dramatically in recent years due to the constantly changing global attack surface. According to reports, the number of malware assaults in 2020 increased by 358% over the year before [22, 26]. Long into 2021, cyberattacks continued to rise significantly, with an estimated 236.1 million ransomware assaults taking place worldwide in the first half of 2022 alone.

According to the World Economic Forum (WEF), this nation would have the third-largest economy in the world if all cybercrime were unified under one flag [1]. In 2022 alone, cybercrime resulted in losses of USD eight trillion. There is no indication that the rise in cybercrime will slow down very soon; on the contrary, it is becoming more and more prevalent.

As mentioned earlier, there were over 236 million ransomware assaults in 2022's first half alone. Malware that encrypts a victim's hard disc and holds it for "ransom" until demands are fulfilled is an apparently recent threat that makes news. Ransomware, however, has been a danger for a while. International attention was originally drawn to the WannaCry ransomware in May 2017 [4, 9]. The enormous effect of WannaCry made it special. The WannaCry assault is thought to have caused billions of dollars in losses, infecting over 300,000 machines across 150 nations. A new trend of cyberterrorism—holding hospitals, schools, and colleges hostage, with little regard for whether the victims pay the ransom or not—was the most catastrophic danger that WannaCry brought about, rather than the possibility of large ransom demands [9, 10]. The main objective of this operation is to severely impair vital infrastructure.

Indeed, the Cybersecurity and Infrastructure Security Agency (CISA) said that "water, hospitals, and K-12"—which in the US refers to kindergarten through the twelfth grade—are their top priorities for 2023. These industries have large attack surfaces, little resources, and are often the target of ransomware. About 290 hospitals in the US were impacted by at least 25 ransomware attacks on "hospitals and multi-hospital health systems" in 2022, according to cybersecurity company Emsisoft. Nearly 1 million students worldwide suffered as a result of 67 ransomware attacks on K–12 schools in 2021 [14, 19]. The outage was estimated to have cost USD 3.5 billion.

In the realm of symmetric cryptography, lightweight primitives are essential for creating safe, quick, and energy-efficient solutions that may be used to high-bandwidth applications as well as embedded devices in resource-constrained settings. The National Institute of Standards and Technology (NIST) started a lightweight cryptography standardisation procedure in 2018 as part of this endeavour [13, 16]. By comparing submitted designs of block cyphers, hashing algorithms, authentication codes for messages (MACs), authenticated encryption with associated data (AEAD), and pseudorandom functions (PRFs), a new standard for lightweight cryptography was to be chosen. The Ascon family was chosen for standardisation on February 7, 2023, marking the conclusion of the process.

PHYSICAL SIDE-CHANNEL ATTACKS

Physical side-channel attacks take use of the secret's information leakage via the device's physical characteristics, such as its power consumption and electromagnetic emissions (EM) [14, 25]. For example, the Differential Power Analysis (DPA) takes use of the intrinsic relationships between the device's power usage and the secret-key dependent data being processed. Since then, it has been shown that several cryptographic implementations are susceptible to these kinds of attacks. As a result, researchers have put forward practical strategies to lessen these threats. However, as cryptography was the sole sector that needed secrecy, the side-channel analysis study was mostly concerned with safeguarding cryptographic implementations. However, side-channel assaults have recently turned ML models into profitable targets as well [5]. Additionally, stealing the model's internals helps adversarial attacks that try to maliciously misclassify the model.

It should come as no surprise that edge-based ML accelerators are vulnerable to physical side-channel attacks, which are simple to implement provided the adversary has physical access to the device. Reverse engineering ML models to a multi-tenant cloud-based FPGA environment is also feasible due to recent advancements in remote physical side-channel assaults [5, 9]. Research on creating effective defences is still in its infancy, however. The development of effective and reliable side-channel defences for ML applications is urgently needed, as the most recent market research projects a massive increase in sales of edge-based ML hardware in the next years.

LIGHTWEIGHT COUNTERMEASURE STRATEGIES AGAINST SIDE-CHANNEL ANALYSIS

We provide the following simple implementation techniques to fend against side-channel attacks:

- **Masking with Periodic Refresh:** In order to randomise the S-Box calculations and eliminate the link between the sensitive data and power leakage during the block cipher's execution, masking of the nonlinear substitution boxes (S-Boxes) is a common countermeasure against DPA [11, 22].The mask should be uniformly selected at random, meaning it should be changed after each encryption, for optimal security. For lightweight applications aimed at devices with limited resources, this is fairly unrealistic due to its high cost in terms of both space overhead and throughput loss [9]. By renewing the mask at pre-specified intervals rather than after each encryption, we propose an alternate design technique that trades efficiency for security. We concentrate specifically on serialised block cypher implementations, a common hardware implementation technique for applications with limited space.

- **Choice of S-Boxes:** A thoughtful selection of S-Box for the block cypher makes the periodic refresh technique much more effective. Specifically, we use the modified transparency order (MTO) measure to examine the resistance of various S-boxes to differential power analysis (DPA). Through simulation experiments, we demonstrate that there is a direct relationship between an S-Box's MTO value [9, 10] and the appropriate refresh rate for a masked implementation of the same in order to provide respectable side-channel security. In other words, a thoughtful S-Box selection informed by the MTO measure may enhance throughput and lower the cost of recurring mask refreshes.
- **Multi-Round Shuffling:** Any key-dependent action in a shuffled implementation is disordered in time, making it harder for the adversary to predict when a particular key-dependent operation will take place [11].

In real-world applications, cryptography must take Side-Channel Attack (SCA) defences into account. SCA has been successfully applied by researchers to classical cryptography's hardware circuit. The periods before and following registration updates are the main targets of their attacks. Due to significant route delays, traditional cryptography is often implemented in loop architecture [18, 19]. This causes circuit designers to employ a large number of registers to store intermediate data, and the clock has a direct impact on the register updates. When registers are changed, they produce significant dynamic power consumption, which is readily intercepted and exploited by adversaries. It is difficult to identify a precise attack position on the power consumption curve since the unrolled architecture does not include clocks or registers. Second, the gathered power consumption curve will have a poor signal-noise ratio (SNR) due to the unrolled architecture's lengthy critical path and several glitch events in between. With the restriction of clearing the datapath after each encryption, the study demonstrates that unrolled DES has some resistance to DPA and CPA. According to the study, a first-order CPA attack may be thwarted by an unrolled MAC-PHOTON.

Because SCAs compute instructions sequentially, they are relatively straightforward to apply to software implementations of cryptographic algorithms like AES. However, because hardware modules compute instructions in parallel, they are challenging to build. Apart from conventional SCAs, several ML and DL models have shown noteworthy significance for power profile SCAs [18]. ML and DL-based power profiling attacks against various FPGAs have been suggested by many researchers.

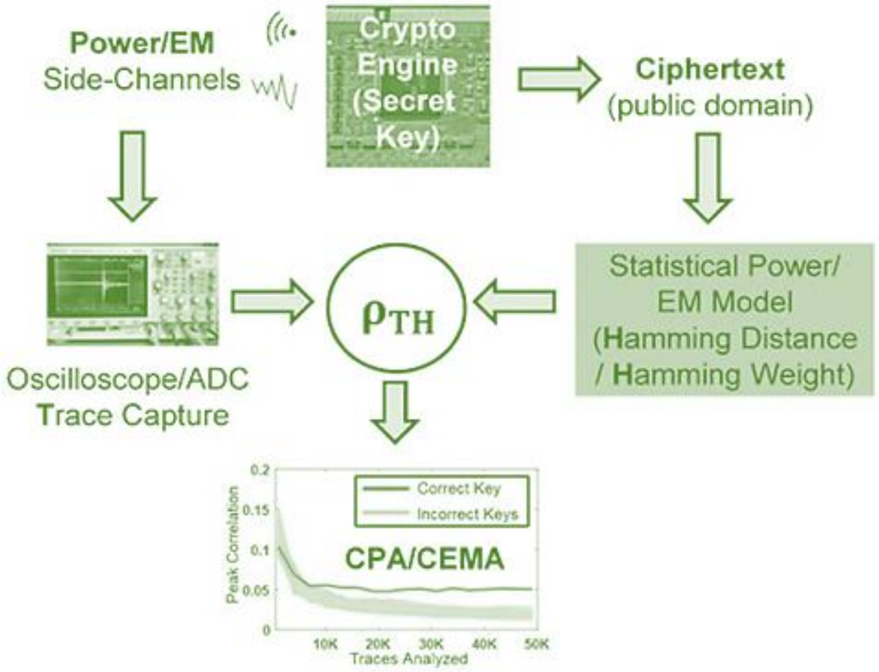

Fig. 2 Power leakage-based SCA types. [19, 23]

LITERATURE REVIEW

Mukhopadhyay, S. (2018) [15] For Internet of Things edge devices, designing an encryption engine that is both secure and very lightweight is a major issue. In order to provide a high degree of side-channel security with minimal performance, power, and area overheads, this paper proposes an optimised data-path architecture for the 128-bit SIMON (SIMON128), a lightweight block cypher. It also investigates the system level design space for an ultralow power image sensor node for secure communication. In order to simultaneously improve energy efficiency and defence against power-based side-channel analysis (PSCA) assaults, many data-path topologies for SIMON are investigated.

Jose, D. (2021) [16] It is believed that security is more crucial in embedded systems and should be assessed every minute. There are many real-time issues with a current encryption system, such as fixed encryption keys and ineffective side-channel attack detection. The following contributions are included in this research work, which addresses the shortcomings of current systems and focusses on integrating strong machine learning algorithms for collecting secret key information employing countermeasures approach using chaotic logistic maps: (a) Creating data sets from the power consumption traces recorded from ARTIX-7 FPGA chips during the Elliptical Curve Cryptography (ECC) operating process (b) Detection and classification of side-channel assaults using single feed-forward learning machines with high speed and accuracy (c) Using 3-D logistic maps for the attacked bits, create chaotic countermeasures.

Cheong, M., (2025) [17] Security has become more important as cloud and embedded systems are increasingly integrated. New risks have continuously surfaced despite attempts to put defences against assaults into place. The most prominent use of deep learning (DL) is in side-channel disassembly attacks, which reveal cloud-to-things functions. This emphasises how important it is to have efficient methods for evaluating how resilient a system is to these kinds of assaults. In order to conceal countermeasures in a fully simulated embedded system, we created a reliable instruction-level side-channel disassembler in this work.

Ahmed, A. A., Hasan, (2024) [18] Significant progress has been made in the use of deep learning for side channel attacks between 2018 and 2024. This compromises the security of implementations of cryptographic algorithms. The goal is to theoretically monitor a semiconductor doing encryption for certain forms of data loss, such as power consumption. Then, using knowledge of the underlying encryption technique, a model is trained to determine the encryption key. After that, the model is applied to traces taken from a victim chip in order to retrieve the encryption key.

Kamrul Hasan, M., (2024) [19] This study investigates a unique approach to SCA profiling in order to improve Deep Learning (DL) models and solve compatibility issues. This study suggests Convolutional Neural Networks as an alternative to defences that concentrate on misalignment. Though it is still acceptable, "Time-Delay Convolutional Neural Networks" (TDCNN) are more accurate than "Convolutional Neural Networks." Indeed, TDCNNs are neural networks that use convolution and are trained on a single piece of spatial input, similar to side-channel tracings. However, considering the current rise in popularity of CNNs, especially after 2012, when the CNN framework (Alex-Net) won the prestigious Image Net Large Scale Visual Recognition Competition, a competition for image detection, a unique TDCNN has been named out in the DL literature.

(Ibraheem, A. S., 2024) [20] One of the main drivers of cloud computing is the use of Internet-based technology in the modern day. Because cloud computing enables cost-effective transmission, storage, and heavy processing, the notion has become very popular. The objective is to reduce an individual's total cost by giving end users access to distant storage and data analysis capabilities via the utilisation of shared computer resources. However, because of security and privacy issues, consumers are still reluctant to embrace this technology.

Usman, O. L. (2024) [21] In order to comprehend the flaws in the implementation of cryptography, researchers are investigating the use of convolutional neural networks (CNNs) in side-channel attacks. CNNs are capable of autonomously learning hierarchical features from power consumption or electromagnetic radiation during cryptographic operations. In order to derive secret keys and extract meaningful representations, researchers train CNNs using side-channel input. CNNs provide a workable paradigm for profiling side-channel analysis assaults, and deep learning methods are useful for assessing the security of embedded systems..

Alzuabidi, I. A. (2024) [22] Machine learning methods are used to both prevent Differential Power Analysis (DPA) attacks and identify the malicious intent behind side-channel attacks on cryptographic systems. Specifically, using the DPA Challenge Dataset, which contains power traces of AES encryption processes, we provide a thorough, step-by-step methodology that includes data collection, pre-processing, feature extraction, and model evaluation. Noise reduction, normalisation, and segmental processing are all part of the pre-processing of the gathered data, from which pertinent characteristics may be extracted using fundamental statistical and frequency domain analysis. After that, Support Vector Machines (SVMs) are trained and evaluated to categorise and then forecast attack scenarios based on the attributes that are later produced.

Kanchana Bhaaskaran, V. S. (2024) [23] Side channel attacks, of which the power analysis attack is the most common, may compromise cypher implementations that provide security for smart grid systems. At the lowest level of abstraction, circuit level safeguards provide security. In this research, we use deep learning architecture to investigate the effectiveness of secure adiabatic logic style-based VLSI implementation against power analysis assaults.

Ahmed, S. (2025) [24] As gadgets become more interconnected, the Internet of Things (IoT) has transformed both everyday life and enterprise. IoT device growth has raised security threats, meanwhile, necessitating strong defences for private information and vital infrastructure. For protecting IoT devices while striking a balance between moderate throughput, low power consumption, and minimum space utilisation, the Advanced Encryption Standard (AES) continues to be the industry standard. Using field programable gate arrays (FPGA) and application-specific integrated circuit (ASIC) implementations, this paper provides a thorough analysis of the most recent lightweight AES architectural designs, including optimisations to the Substitution Box (S-Box), Sub-Bytes, Shift Rows, Mix Columns, and Add Round Key steps. It evaluates the effects of these changes on gate count, area, maximum frequency, power consumption, and throughput.

In previous studies, all of the researchers focus on whether the attack is successful or not. No one focused on the required training time $\boldsymbol{T_t}$ for the model, the secret key prediction time $\boldsymbol{T_p}$**,** and the time needed for a successful attack $\boldsymbol{T_a}$**.** Time is a crucial consideration when doing security research on hardware devices, as is well known [26]. We have shown several tests and decreased the amount of power traces needed to attack [28]. We have examined every facet of time in this study and verified that ML/DL models are capable of carrying out a successful SCA. The following are our main contributions:

- We have provided a concise synopsis of techniques for FPGA power analysis.
- Additionally, we have provided a comparison between our proprietary CNN model and cutting-edge models such as CNN VGG-16, long short-term memory (LSTM), and MLP.
- We have presented a novel methodology that describes the required $\boldsymbol{T_t}$ for the model, $\boldsymbol{T_p}$**,** and $\boldsymbol{T_a}$**.**

METHODOLOGY

Dataset

We utilised the AES-HD (Advance Encryption System Hamming Distance) dataset of FPGA power traces, which is freely accessible at AES-HD GitHub. An unsecured AES128 module provided the AES-HD dataset. VHDL (VHSIC Hardware Description Language) was used to write the round-based design of AES, which requires 11 clock cycles for a single encryption [26]. UART was used to connect the AES module to the external circuit. The Xilinx Vertix-5 FPGA of a SASEBO GII board was used to build the design, which had 747 flip-flops and 1850 LUT (lookup tables).

Using a high-sensitivity near-field electromagnetic probe over the power line's decoupling capacitor, the Teledyne LeCroy Waverunner 610zi oscilloscope recorded the power traces. In the last AES encryption round, the power side channel leakage was assessed using the Hamming Distance (HD) model. A signal-to-noise-ratio (SNR) of 0.0096 indicated that the power traces were noisy [27]. There were 1250 characteristics in each of the 100,000 power traces that were gathered against the 100,000 randomly generated plain texts. Since, to the best of our knowledge, this is the only recent FPGA power dataset that is publicly accessible, we utilised it to apply our suggested technique.

Proposed Framework

The work described in [28] serves as the foundation for our suggested framework, which is used to finalise the block-level representation in Fig. 3.

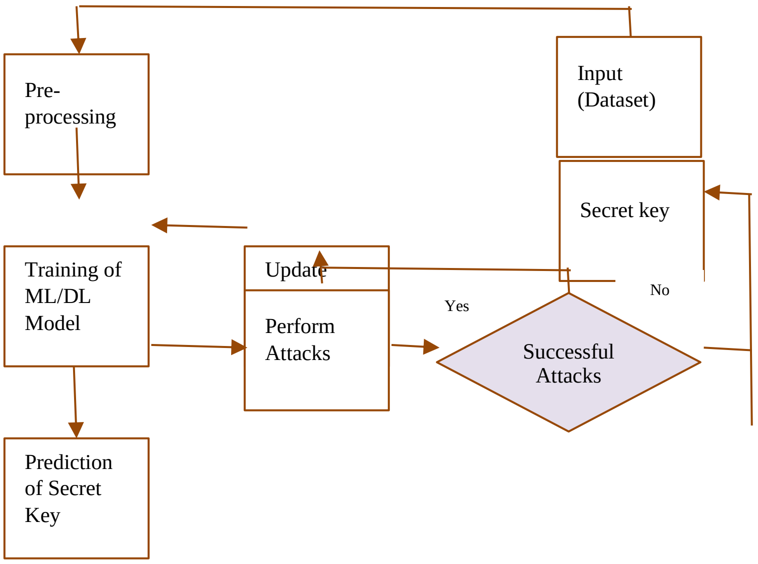

Fig. 3 Proposed Method. [28]

- **Pre-processing:** We employed the most used methods, standardisation and normalisation, to scale features in pre-processing [28, 29]. Equation 1 illustrates how the standardisation approach uses the mean and standard deviation to modify the data.

$$X_{Std} = \frac{X-Mean}{Std}. \ldots\ldots\ldots\ldots\ldots\ldots 1$$

Data in the range of [0,1] or [-1,1], as shown in equation 2, is transformed using the normalisation procedure. These methods facilitate quicker ML/DL model convergence and training.

- **Training of ML/DL Models:** Only 75,000 randomly selected power connections out of all the power traces were utilised. Power traces were split into two sets: test/attack (25000 for testing) and train (45000 for training and 5000 for validation). Four distinct ML/DL models (custom CNN, CNN VGG-16-Like, MLP, and LSTM) were applied to the training set after data pre-processing.

$$X\_Norm = \frac{X-X_{Min}}{X_{Max}-X_{Min}} \ldots\ldots\ldots\ldots\ldots\ldots 2$$

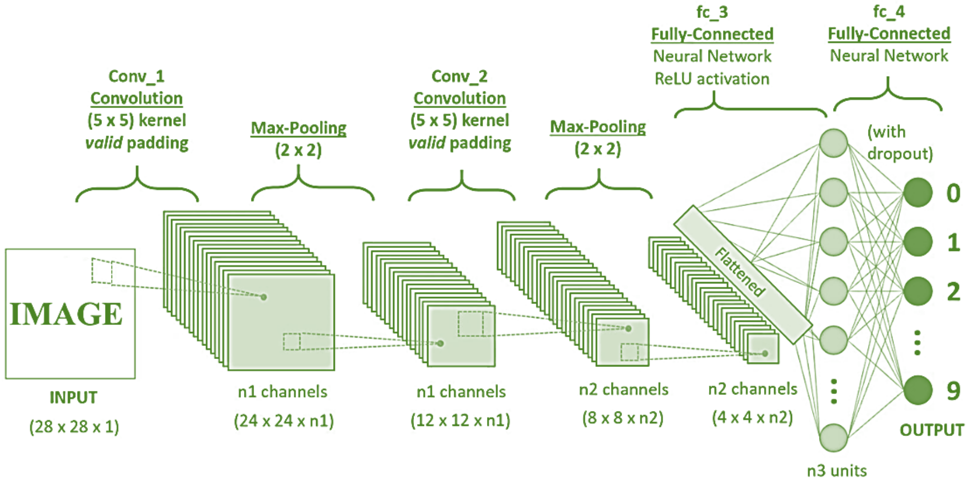

Fig. 4 The Structure of Our Personalised CNN Model. [25, 26]

As seen in Fig. 4, our optimal custom CNN model included an input layer, a convolution layer, a batch normalisation layer, an average pooling layer, a flattens layer, a fully connected layer, and an output layer. The input layer's size was 1x1250 as we made use of every feature in the power trace. 1-D convolution was carried out in the convolution layer.

In the output layer, we divided our output into 256 classes as the AES S-box value ranges from 0 to 255. Fig. 5 displays our best custom CNN model's block diagram.

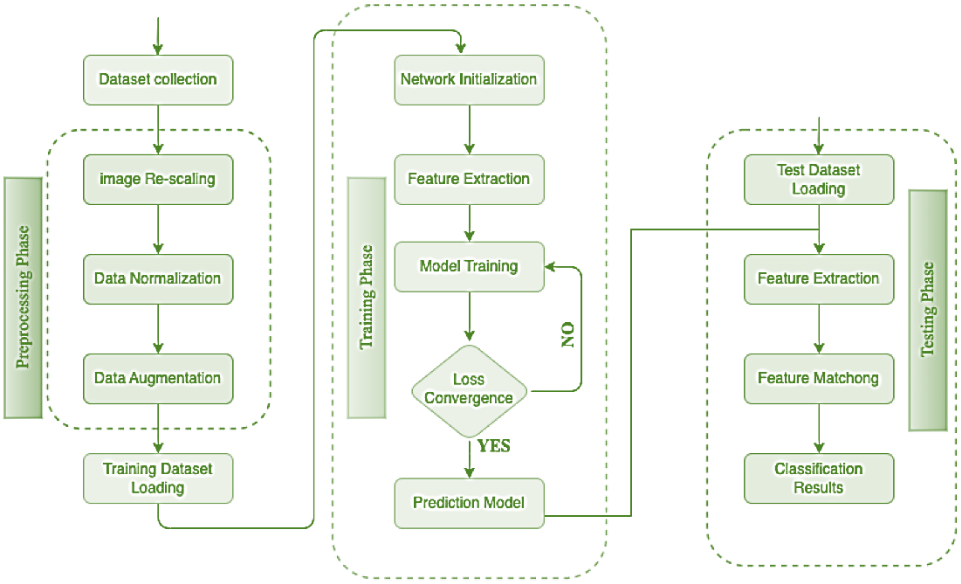

Fig. 5 Block diagram of our custom CNN model. [29]

- **Prediction of Secret Key:** In the prediction stage, the model labels each test data set with a number between 0 and 255 and uses the trained model to predict a secret key. Following prediction completion, we received a 16-byte projected key that is used throughout the attack phase.
- **Profiling Attack Model:** We performed an attack against AES using a power-profiled assault after predicting the secret key. We varied the attack stage settings (number of attacks, attack ciphertext, attack power traces, and predicted key from the previous prediction stage) in order to achieve a successful attack [22, 28]. Although we employed 100 and 200 attacks consistently across all studies, the 100 assaults had higher power profiles and took less time to execute. With the use of our ML/DL models, we created a profile or template for this attack utilising power traces and the cryptographic algorithm's known actual key. In the end, we applied an attack on the power traces of the test set. To determine the degree to which the attack power traces and critical information are associated, our models performed template matching in this stage.

RESULT AND DISCUSSION

The experimental setup, test examples, [26,28], and comparison of our suggested model with the most advanced models are all included. Guessing entropy (GE) is a special performance statistic needed to solve the research challenge.

Performance Metric

By lowering the quantity of attack power traces, the attacker hopes to find a prediction model that improves the secret key retrieval during the attack stage. GE, or success rate, is a statistic used in side-channel attacks that characterises the average entropy/rank of the actual secret key among all of the major hypotheses. We assume that the secret key is related to the guessing value. We define a successful side-channel assault as one in which the GE consistently equals 1. This GE of the correct secret key tells us how successfully our ML/DL models are recovering the secret key. In our paper, GE plays a crucial role in our ML/DL model studies and demonstrates how effective these models are in producing outcomes.

Evaluation of Experiments and Results

We have used a variety of filter sizes, activation functions, epochs, layers, and optimisers, as we covered in B2. The number of attack power traces where GE equalled one, the time needed to train a model, the time needed to forecast the secret key, and the time needed for the attack were the four essential criteria we took into account when evaluating the outcomes. Accordingly, as shown in Table 1, we have separated these trials and their assessment into three distinct situations (assessment Case 1, Evaluation Case 2, and Evaluation Case 3) based on three distinct optimisers (Adam, Nadam, and Ada-Delta). All of the tests were conducted and assessed using Colab GPU in Google Colab.

Table 1 An overview of evaluation situations together with the corresponding parameters.

Evaluation Cases	**Optimizer**	**Activation Function**	**No of Filters**	**No. of Layers**	**No. of Epochs**
Cases 1	AdaDelta	Relu&Selu	2, 4, 8 & 16	8, 10, 12, & 14	20, 50, 100, & 200
Cases 2	Nadam	Relu&Selu	2, 4, 8 & 16	8, 10, 12, & 14	20, 50, 100, & 200
Cases 3	Adam	Relu&Selu	2, 4, 8 & 16	8, 10, 12, & 14	20, 50, 100, & 200

We used the Adam optimiser and learning rate = 0.00001 for both models in evaluation case 1 from Table 1. In order to display the GE assessment graphs for complete key recovery, we conducted 24 distinct tests on all models. In experiment 1, we used selu as an activation function and 20 epochs.

We repeated experiment 1 with 4, 8, and 16 filters and 10, 12, and 14 layers to get the best and worst results. We used 20 epochs with relu as an activation function in experiment 2.

Similar to experiment 1, the remainder of the experiment was identical:

We used 50 epochs with selu as an activation function in experiment 3. We used relu as an activation function and 50 epochs in experiment 4. We used selu as an activation function and 100 epochs in experiment 5. We used relu as an activation function and 100 epochs in experiment 6. We used selu as an activation function and 200 epochs in experiment 7. In experiment 8, we used relu as an activation function and 200 epochs. Similar tests were conducted for assessment instances 2 and 3, with the exception of the optimisers.

The selu activation function, 20 epochs, and Nadam optimiser produced outstanding outcomes for our customised CNN model. The key begins to unveil when the GE equals 25 after 130 assault power traces.

As shown in Fig. 6, our customised CNN model with two filters used around 570 attack power with *Tt=46.18 seconds*, *Tp=2.58 seconds*, and *Ta=118.96 seconds* to recover the whole secret key.

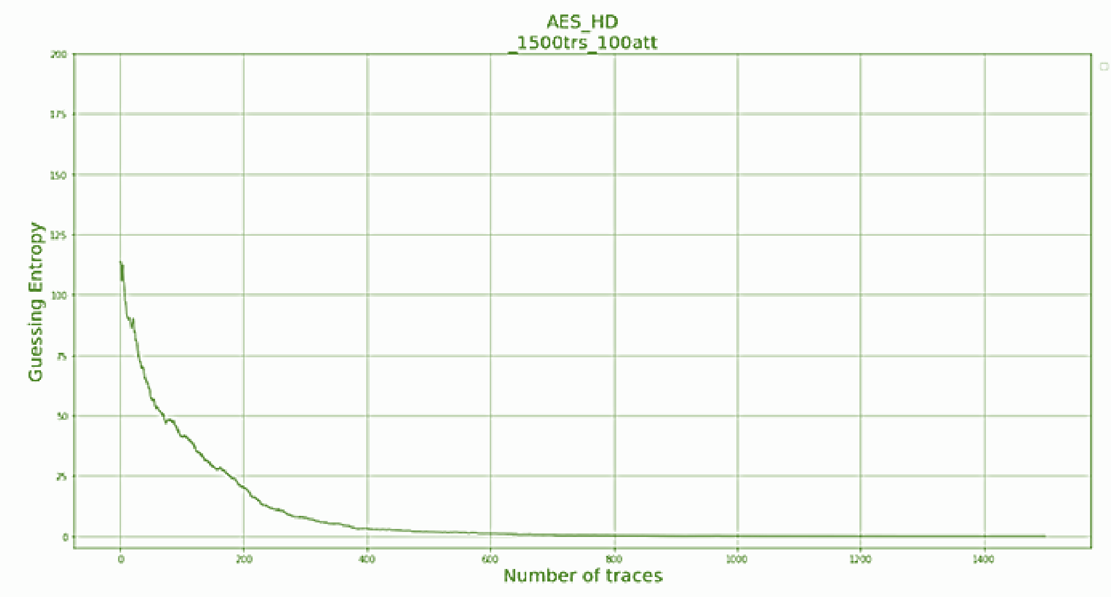

Fig. 6 GE's best performance using our customised CNN model.

The MLP model's second-best results were obtained using the Adam optimiser, 20 epochs, and relu activation function.

As shown in Fig. 7, the MLP model with 12 layers used around 3200 power traces with *Tt=97.89seconds*, *Tp=5.49seconds*, and *Ta=138.97seconds*.

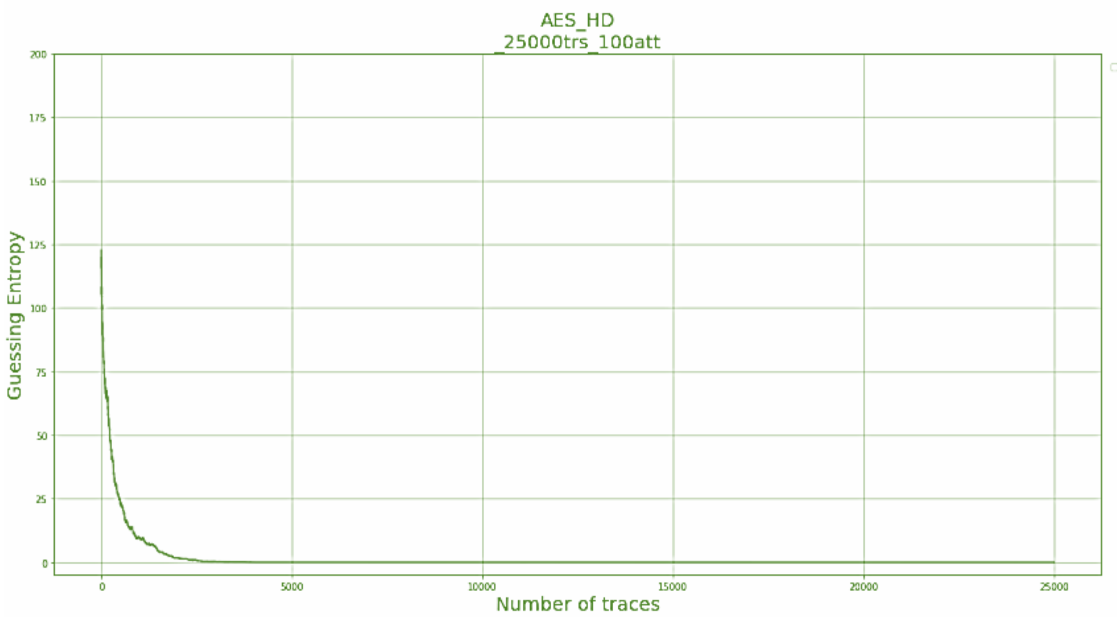

Fig. 7 GE's second-best performance with our customised CNN model.

Discussion

We briefly reviewed the findings and assessment techniques of the earlier state-of-the-art studies before discussing our findings. 5000 power traces were needed to uncover all 16 bytes of the secret key from the cryptographic algorithm AES using random forest (RF) and support vector machines (SVM) [29, 30]. To balance the data, they employed the synthetic minority oversampling technique (SMOTE), and for accuracy, they used GE.

More than 2000 power traces were needed to extract a single byte of the secret key when they used the CNN model. They used the GE method to evaluate their successful power profiled attack [31]. They demonstrated the LSTM auto-encoder model, which required almost 3700 power traces to recover the secret key's 16 bytes. They used the key rank at which the secret key was disclosed to assess their attacks.

They used GE to assess their model's performance [31, 32]. Using 3700 traces, the CNN model was constructed and all 16 bytes of the secret key were made public. By using the CNN model, they were able to uncover all 16 bytes of the secret key using only 831 power traces [28, 29]. To assess their findings, they used the GE approach. 1050 power traces were needed in their model in order to successfully launch an attack and expose every private key. They assessed their performance using GE as a measure. To uncover all 16 bytes of the secret key, 2100 power traces were required. They assessed their model's performance using the average rank approach [33].

As we proceed with our experiments and findings, we determined that the nadam optimiser, selu activation function, two filters, and twenty epochs are the optimal parameters of our best custom CNN approach. Additionally, we investigated how our suggested custom CNN model demonstrated an incredible and successful attack with a manageable training time and needed almost half as many attack power traces as the base model. The MLP model was the second-best model in our experiments and also demonstrated noteworthy results when it came to the power-profiled SCAs [34, 35].

We discovered after a lot of testing that the ideal parameters for the MLP model are 20 epochs, 12 layers, relu activation function, and Adam optimiser. This dataset was too difficult for the CNN VGG-16-like and LSTM models to successfully attack [36]. Table 3 presents a comprehensive overview of power-profiled SCAs using ML and DL models.

Table 3 ML and DL-Based Power Profiled Side-Channel Attack Synopsis.

Works	**Target Devices**	**Classifier**	**Traces of Requirements for Successful Attacks**	**Revealing Bytes**	**Method of Evaluations**
14	Virtex-5FPGA	RF &SVM	5000	16	Key rank
15	Virtex-5FPGA	CNN	800	Single	PI
16	Virtex-5FPGA	LSTM	<2000	16	GE
17	Antux-7 FPGA	CNN VGG-16 Like	3600	Single	PI

18	ASIC	LSTM	1500	16	PI
19	Virtex-5FPGA	CNN VGG-16 Like	24900	Single	PI
20	Spartan-6 FPGA	CNN	3900	16	Key Rank
21	Virtex-5FPGA	CNN	879	Single	GE
22	Virtex-5FPGA	CNN VGG-16 Like	1098	16	Accuracy
23	Virtex-5FPGA	CNN	2109	16	Key Rank
Proposed	Virtex-5FPGA	CNN VGG-16 Like	<25000	16	GE
Proposed	Virtex-5FPGA	MLP	3600	16	PI
Proposed	Virtex-5FPGA	CNN	489	16	Key Rank

CONCLUSION

We provide a thorough examination and analysis of the primary features and contributions to power analysis side channel attacks on FPGA as a conclusion to this work. First, we provide a brief explanation of how to perform an FPGA power analysis. Two methods of FPGA power analysis have been identified based on prior state-of-the-art research. There are two methods for measuring power: physically using an oscilloscope, UART, etc., and remotely using sensors like ADC, LDC, RO, etc.

Second, we provide a quick explanation of SCAs and their many forms. Two main categories of power analysis-based SCAs—TA, FIA, MLA, DLA, SPA, DPA, and CPA attacks—are profiled and non-profiled. Thirdly, we evaluate the power analysis-based SCAs using machine learning and deep learning methods (MLP and CNN).

On the AESHD dataset, our suggested custom CNN model demonstrates an amazing and effective assault that uses about half as many attacks power traces (around 570) with very short *Tt=46.18 seconds*, *Tp=2.58 seconds*, and *Ta=118.96 seconds* of time.

With tiny *Tt=97.89seconds*, *Tp=5.49seconds*, and *Ta=138.97seconds* time, our MLP model demonstrated a complete key after 3200 power traces, demonstrating exceptional performance on this noisy dataset.

Our suggested architecture is a model that is lightweight, simple to train, and requires few parameters to successfully execute power analysis attacks on FPGA.

Future research will primarily focus on real-time execution of power profiled assaults on FPGA and our own collection of power traces.

REFERENCES

[1] Sagu, A.; Gill, N.S.; Gulia, P.; Singh, P.K.; Hong, W.C. Design of Metaheuristic Optimization Algorithms for Deep Learning Model for Secure IoT Environment. Sustainability 2023, 15, 2204.

[2] Eskandari, M.; Janjua, Z.H.; Vecchio, M.; Antonelli, F. Passban IDS: An Intelligent Anomaly-Based Intrusion Detection System for IoT Edge Devices. IEEE Internet Things J. 2020, 7, 6882–6897.

[3] Almiani, M.; AbuGhazleh, A.; Al-Rahayfeh, A.; Atiewi, S.; Razaque, A. Deep recurrent neural network for IoT intrusion detection system. Simul. Model. Pract. Theory 2020, 101, 102031, Modeling and Simulation of Fog Computing.

[4] Tavallaee, M.; Bagheri, E.; Lu, W.; Ghorbani, A.A. A detailed analysis of the KDD CUP 99 data set. In Proceedings of the 2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications, Ottawa, ON, Canada, 8–10 July 2009; pp. 1–6.

[5] Saba, T.; Rehman, A.; Sadad, T.; Kolivand, H.; Bahaj, S.A. Anomaly-based intrusion detection system for IoT networks through deep learning model. Comput. Electr. Eng. 2022, 99, 107810.

[6] Ngo, D.M.; Lightbody, D.; Temko, A.; Pham-Quoc, C.; Tran, N.T.; Murphy, C.C.; Popovici, E. HH-NIDS: Heterogeneous Hardware-Based Network Intrusion Detection Framework for IoT Security. Future Internet 2023, 15, 9.

[7] Dhananjay, K.; Salman, E. Charge Based Power Side-Channel Attack Methodology for an Adiabatic Cipher. Electronics 2021, 10, 1438.

[8] Morales Romero, J.d.J.; Reyes Barranca, M.A.; Tinoco Varela, D.; Flores Nava, L.M.; Espinosa Garcia, E.R. SCA-Safe Implementation of Modified SaMAL2R Algorithm in FPGA. Micromachines 2022, 13, 1872.

[9] Mangard, S.; Oswald, E.; Popp, T. Power Analysis Attacks: Revealing the Secrets of Smart Cards; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2008; Volume 31.

[10] Zhou, F.; Zhang, B.; Wu, N.; Bu, X. The design of compact SM4 encryption and decryption circuits that are resistant to bypass attack. Electronics 2020, 9, 1102.

[11] Bhasin, S.; Guilley, S.; Sauvage, L.; Danger, J.L. Unrolling cryptographic circuits: A simple countermeasure against side-channel attacks. In Proceedings of the Cryptographers' Track at the RSA Conference, San Francisco, CA, USA, 1–5 March 2010; Springer: Berlin/Heidelberg, Germany, 2010; pp. 195–207.

[12] Pub, F. Data encryption standard (des). In FIPS PUB; NIPS: Gaithersburg, MD, USA, 1999; pp. 46–583.

[13] Kocher, P.; Jaffe, J.; Jun, B. Differential power analysis. In Proceedings of the Annual International Cryptology Conference, Santa Barbara, CA, USA, 15–19 August 1999; Springer: Berlin/Heidelberg, Germany, 1999; pp. 388–397.

[14] Nalla Anandakumar, N. SCA Resistance Analysis on FPGA Implementations of Sponge Based MAC-PHOTON. In Proceedings of the International Conference for Information Technology and Communications, Bucharest, Romania, 11–12 June 2015; Springer: Berlin/Heidelberg, Germany, 2015; pp. 69–86.

[15] Singh, A., Chawla, N., Ko, J. H., Kar, M., & Mukhopadhyay, S. (2018). Energy efficient and side-channel secure cryptographic hardware for IoT-edge nodes. IEEE Internet of Things Journal, 6(1), 421-434.

[16] Illuri, B., & Jose, D. (2021). RETRACTED ARTICLE: Design and implementation of hybrid integration of cognitive learning and chaotic countermeasures for side channel attacks. Journal of Ambient Intelligence and Humanized Computing, 12(5), 5427-5441.

[17] Alabdulwahab, S., Cheong, M., Seo, A., Kim, Y. T., & Son, Y. (2025). Enhancing deep learning-based side-channel analysis using feature engineering in a fully simulated IoT system. Expert Systems with Applications, 266, 126079.

[18] Ahmed, A. A., Hasan, M. K., Aman, A. H., Safie, N., Islam, S., Ahmed, F. R. A., ... & Rzayeva, L. (2024). Review on hybrid deep learning models for enhancing encryption techniques against side channel attacks. IEEE Access.

[19] Abbas Ahmed, A., Kamrul Hasan, M., Azman Mohd Noah, S., & Hafizah Aman, A. (2024). Design of time-delay convolutional neural networks (TDCNN) model for feature extraction for side-channel attacks. International Journal of Computing and Digital Systems, 16(1), 341-351.

[20] Younus, Y. M., Ibraheem, A. S., & Tuama, M. H. (2024). Improved Machine Learning Techniques for Precise DoS Attack Forecasting in Cloud Security. (Humanities, social and applied sciences) Misan Journal of Academic Studies, 23(52), 122-132.

[21] Fattah, A., Mutashar, H. J., & Usman, O. L. (2024). Design of Deep Learning Techniques for Side-Channel Attacks on Masked 128-bit AES Implementations. AlKadhim Journal for Computer Science, 2(1), 86-96.

[22] Alzuabidi, I. A. (2024). Application of machine learning techniques for countering side-channel attacks in cryptographic systems. Alkadhim J. Comput. Sci, 2(3).

[23] Banu, A. J., Prathiba, A., Shyam Krishna, S., Peddhibhotla, S., & Kanchana Bhaaskaran, V. S. (2024). Profiled Side Channel Power Attack on Charge Balancing Symmetric Pre-Resolve Adiabatic Logic PRESENT S-Box Using Convolutional Neural Networks. Smart Grids as Cyber Physical Systems: Smart Grids Paving the Way to Smart Cities, 2, 245-275.

[24] Ahmed, S., Ahmad, N., Shah, N. A., Abro, G. E. M., Wijayanto, A., Hirsi, A., & Altaf, A. R. (2025). Lightweight AES Design for IoT Applications: Optimizations in FPGA and ASIC with DFA Countermeasure Strategies. IEEE Access.

[25] G. Zaid, L. Bossuet, A. Habrard, and A. Venelli, "Methodology for efficient cnn architectures in profiling attacks," IACR Transactions on Cryptographic Hardware and Embedded Systems, vol. 2020, no. 1, pp. 1–36, 2020.

[26] M. Jin, M. Zheng, H. Hu, and N. Yu, "An enhanced convolutional neural network in side-channel attacks and its visualization," arXiv preprint arXiv:2009.08898, 2020.

[27] T. Kubota, K. Yoshida, M. Shiozaki, and T. Fujino, "Deep learning side-channel attack against hardware implementations of aes," Microprocessors and Microsystems, vol. 87, p. 103383, 2021.

[28] J. Kim, S. Picek, A. Heuser, S. Bhasin, and A. Hanjalic, "Make some noise. unleashing the power of convolutional neural networks for profiled side-channel analysis," IACR Transactions on Cryptographic Hardware and Embedded Systems, pp. 148–179, 2019.

[29] A. Al Arafat, Z. Guo, and A. Awad, ''Vr-spy: A side-channel attack on virtual key-logging in vr headsets,'' in Proc. IEEE Virtual Reality 3D User Interfaces, Mar. 2021, pp. 564–572,

[30] B. Colombier, V.-F. Dragoi, P.-L. Cayrel, and V. Grosso, ''Messagerecovery profiled side-channel attack on the classic McEliece cryptosystem,'' IACR Cryptol. ePrint Arch., vol. 1, pp. 1–24, Nov. 2022.

[31] D. Das, J. Danial, A. Golder, S. Ghosh, A. R. Wdhury, and S. Sen, ''Deep learning side-channel attack resilient AES-256 using current domain signature attenuation in 65nm CMOS,'' in Proc. IEEE Custom Integr. Circuits Conf. (CICC), Mar. 2020, pp. 1–4.

[32] K. E. Narayana and K. Jayashree, ‘‘Survey on cross virtual machine side channel attack detection and properties of cloud computing as sustainable material,’’ Mater. Today, Proc.,

[33] Ghazal, Taher M., et al., "Private blockchain-based encryption framework using computational intelligence approach." Egyptian Informatics Journal 23.4 (2022): 69-75.

[34] D. Kwon, Hong, et. al., "Optimizing Implementations of Non-Profiled Deep Learning-Based Side-Channel Attacks," in IEEE Access, vol. 10, pp. 5957-5967, 2022.

[35] Illuri, B., Jose, RETRACTED ARTICLE: Design and implementation of hybrid integration of cognitive learning and chaotic countermeasures for side channel attacks. J Ambient Intell Human Comput 12, 5427– 5441 (2021).

[36] C. Wang, Dani, et. al., "TripletPower: Deep-Learning Side-Channel Attacks over Few Traces," 2023 IEEE International Symposium on Hardware Oriented Security and Trust (HOST), San Jose, CA, USA, 2023, pp. 167-178.