Indonesian Journal of Electrical Engineering and Computer Science
Vol. 42, No. 3, June 2026, pp. 846~855
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v42.i3.pp846-855

Deep learning in cryptanalysis a comprehensive review of techniques, applications, challenges, and future trajectories

Oussama Noui[1], Amine Barkat[2]
[1]Fundamental Teaching Department, Faculty of Sciences of Matter, University of Batna 1, Batna, Algeria
[2]LRE3S2 Laboratory, Higher National School of Renewable Energies, Environment and Sustainable Development, Djerma, Algeria

Article Info

Article history:

Received Aug 20, 2025
Revised Mar 7, 2026
Accepted May 26, 2026

Keywords:

Adversarial machine learning
Convolutional neural networks
Cryptanalysis
Deep learning
Explainable AI
Lightweight cryptography
Machine learning
Side-channel analysis

ABSTRACT

The integration of deep learning (DL) into cryptanalysis represents a paradigm shift, challenging traditional mathematical approaches and unlocking novel attack vectors against cryptographic algorithms and implementations. This comprehensive review synthesizes findings from recent pivotal publications to provide an in-depth analysis of the current state-of-the-art, methodologies, empirical successes, fundamental limitations, and future potential of DL in cryptanalysis. We focus extensively on two primary domains DL-based side-channel analysis (DL-SCA) and DL-enhanced cryptanalysis of symmetric primitives. The review meticulously examines advancements in attack efficiency (reducing the number of traces/queries), robustness against sophisticated countermeasures, automated feature extraction, and the nascent exploration of theoretical foundations. While DL demonstrates remarkable capabilities in automating complex pattern recognition critical to cryptanalysis, significant challenges persist, including the "black-box" nature of models, data dependency, scalability to full cryptographic primitives, and the critical need for explainability and theoretical grounding. This review serves as a foundational resource for researchers and practitioners navigating this rapidly evolving intersection of artificial intelligence and cryptography.

Corresponding Author:

Oussama Noui
Fundamental Teaching Department, Faculty of Sciences of Matter
University Of Batna 1
Batna, Algeria
Email oussama.noui@univ-batna.dz.

1. INTRODUCTION

Cryptography forms the foundation of modern digital security, enabling the protection of data confidentiality, integrity, and authenticity. Cryptanalysis-the adversarial counterpart—aims to uncover weaknesses in cryptographic algorithms or their physical implementations, exploiting either design flaws or vulnerabilities arising from side-channel leakage and fault injection [1]–[4]. Traditional cryptanalysis has long relied on advanced mathematical methods, including differential, linear, and integral techniques. However, the rapid increase in computational capabilities and the emergence of machine learning (ML), particularly deep learning (DL), have introduced a powerful data-driven paradigm for cryptanalytic research [5]–[7].

DL models, especially convolutional neural networks (CNNs), are adept at detecting complex, nonlinear patterns in high-dimensional, noisy datasets-a capability directly relevant to numerous cryptanalytic challenges [8]–[10]. In side-channel analysis (SCA), DL offers several distinct advantages over classical approaches,

- Automated feature extraction DL models can learn discriminative features directly from raw inputs (e.g., power traces, electromagnetic emanations, ciphertext pairs), eliminating the need for manual and often error-prone feature engineering required by traditional SCA or statistical distinguishers [7], [11]–[13].
- Noise and misalignment tolerance CNN-based approaches demonstrate robustness against noise and trace desynchronization, effectively modeling complex dependencies that are difficult to capture analytically [13]–[15].
- Countermeasure resilience techniques such as data augmentation enable DL-based SCA to mitigate the effects of common countermeasures, including random delays, shuffling, and masking [15]–[17].
- Improved attack efficiency in many cases, DL-SCA achieves successful key recovery with significantly fewer traces compared to template attacks or correlation power analysis (CPA). Similarly, in algorithmic cryptanalysis, DL can uncover more efficient distinguishers and differential characteristics than traditional heuristic search methods [18]–[20].

This review synthesizes findings from recent studies to provide an integrated perspective on the role of DL in modern cryptanalysis. Specifically, it examines,

- fundamental concepts and DL architectures employed in cryptanalysis (sections 2–3),
- the transformative impact of DL on side-channel attacks (section 4),
- applications to the cryptanalysis of symmetric cryptographic primitives, particularly block ciphers (section 5),
- limitations, challenges, and the "black-box" interpretability problem (section 6), and
- promising future research directions (section 7).

Unlike existing surveys that primarily focus on either physical side-channel attacks or isolated neural distinguishers, this review provides a unified analysis of DL-based SCA and DL-assisted algorithmic cryptanalysis of symmetric primitives. In addition to systematizing architectures, attack models, and evaluation metrics, this work emphasizes explainability, scalability, and generalization, which are often underexplored. Furthermore, the review offers comparative taxonomies and synthesized tables that highlight practical capabilities, limitations, and open challenges.

This paper follows the IMRaDC structure. Section 1 introduces the background and motivation. Section 2 presents foundational concepts and the review methodology. Sections 3 and 4 describe the methods and architectures employed in DL-based cryptanalysis, with a focus on SCA. Sections 5 and 6 provide results-oriented discussion and critical analysis of DL applications to symmetric cryptanalysis, including limitations and challenges. Section 7 outlines future research directions, and section 8 concludes the paper.

2. FOUNDATIONAL CONCEPTS

2.1. Cryptanalysis primer

Cryptographic goals confidentiality, integrity, authenticity. Attack models ciphertext-only, known-plaintext, chosen-plaintext, chosen-ciphertext. Implementation attacks vs. algorithmic attacks distinguishing attacks targeting physical leakage (SCA, fault attacks) from those targeting the mathematical structure of the cipher. Security metrics success rate (SR), guessing entropy (GE), number of traces to disclosure (NTD) for SCA; probability of differential/linear characteristics, data/time complexity for algorithmic attacks [21]–[23].

2.2. ML fundamentals

Supervised learning learning a mapping from inputs (traces, ciphertexts) to labels (key bytes, distinguishing labels). Dominates cryptanalysis applications [15]. Profiling vs. Non-Profiling Profiling (Supervised) attacks require a controlled profiling phase using a clone device; non-Profiling attacks attempt key recovery directly. Core concepts training/validation/test sets, overfitting/underfitting, loss functions (categorical cross-entropy common for classification), optimization algorithms (SGD, Adam) [24]. General DL workflows, architectures, and application paradigms are discussed in detail in [25].

2.3. DL specifics

Artificial neural networks (ANNs) composed of layers of interconnected neurons (nodes). The depth defines "deep" learning. Multi-layer perceptrons (MLPs) Fully connected networks. Early use in SCA and block cipher distinguishers [26]. CNNs Specialized for grid-like data (images, time-series like SCA traces). Use convolutional layers to extract local features, pooling layers for downsampling. State-of-the-art for DL-SCA [8]. Activation functions rectified linear unit (ReLU) most common, enabling non-linearity. Backpropagation Algorithm for calculating gradients and updating weights during training. D ata Augmentation Artificially expanding the training dataset by applying transformations (e.g., random shifts, jitter, noise addition) to improve robustness and generalization [27].

2.4. Review methodology workflow

This review follows a structured scoping review methodology designed to systematically identify, screen, and synthesize recent advances in DL-based cryptanalysis. The selection process was organized into four stages-identification, screening, eligibility, and synthesis-to ensure transparency and reproducibility.

- Identification: candidate studies were retrieved via keyword-based searches across major digital libraries and cryptography repositories.
- Screening: duplicate records were removed and works not directly related to cryptanalysis or SCA (e.g., general ML papers without cryptographic focus) were excluded.
- Eligibility: included studies applied DL (or advanced ML) to cryptanalysis or SCA and provided empirical evaluation (e.g., GE, SR, NTD, or distinguishing accuracy/advantage).
- Synthesis: selected works were categorized by attack model, target primitive or device, neural architecture, dataset/trace characteristics, and evaluation metrics, enabling a comparative synthesis of capabilities, limitations, and open research challenges.

To clarify the scope and organization of this survey, Figure 1 summarizes the main categories of cryptographic vulnerabilities and the corresponding attack families considered in the literature. Vulnerabilities can be grouped into, i) implementation-level weaknesses, where physical leakage or fault sensitivity enables side-channel and fault-injection attacks; ii) algorithmic-level weaknesses, where statistical biases enable distinguishers and key-recovery methods such as differential and linear cryptanalysis; and iii) protocol/system-level weaknesses related to composition, misuse, or system side effects. ML and DL act as cross-cutting enablers across these categories: in DL-SCA, models primarily support denoising, alignment, and automated feature extraction to enable robust profiling under countermeasures; in algorithmic cryptanalysis, neural distinguishers and learned biases can be combined with classical techniques to guide or filter search; and at the system level, learning-based pattern analysis is occasionally used for anomaly or traffic/timing characterization (though it is less central to the present survey). This taxonomy supports a structured synthesis and motivates the comparative analysis developed in the subsequent sections. In practice, system-level exposure can also stem from misconfiguration of hardware security modules and permission controls [28].

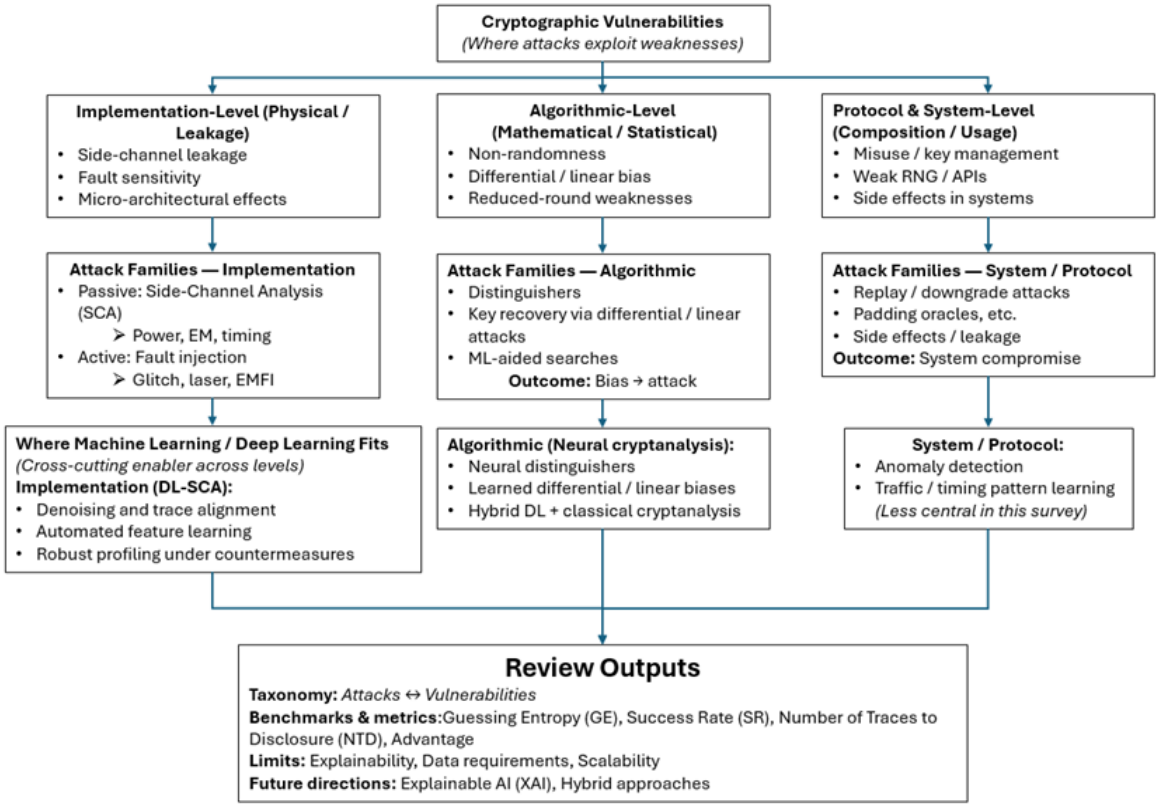

Figure 1. Taxonomy of cryptographic vulnerabilities and corresponding attack families, highlighting where ML and DL are integrated across implementation-, algorithmic-, and protocol/system-level attacks

3. DL ARCHITECTURES FOR CRYPTANALYSIS

3.1. Architectures for side-channel analysis

CNN architectures predominant architecture. VGG-inspired networks (stacked small convolutional layers) are highly effective. Residual connections (ResNet) help train deeper models. Custom architectures tailored to trace characteristics are common [29]. Input representation one-dimension CNNs for single-trace attacks (treating trace as 1D signal). 2D CNNs are sometimes used if traces are pre-processed into spectrograms or combined with other data dimensions. Output layer typically, a softmax layer for classification (predicting a specific key byte value). Regression (predicting a leakage value) is less common but explored. Hyperparameter tuning critical for performance. Includes number/filter size of convolutional layers, pooling strategies, dense layer size, learning rate, batch size. Automated methods (grid search, Bayesian optimization) are increasingly used [30].

3.2. Architectures for algorithmic cryptanalysis

Distinguisher networks often MLPs or CNNs. Input is pairs of plaintexts/ciphertexts or differentials. Output is a probability score indicating whether the input belongs to the cipher or a random permutation. Gohr's Speck attack [11] used a sophisticated CNN.

Key recovery networks less common and harder. Can involve building neural distinguishers for key-dependent differentials or directly predicting key bits (extremely challenging for full ciphers) [5]. Often combined with traditional key search or ranking. Hybrid approaches combining neural distinguishers with classical cryptanalytic techniques (e.g., using a neural distinguisher to filter promising differential paths for traditional key recovery).

3.3. Training methodologies and challenges

Data requirements massive datasets needed, especially for algorithmic cryptanalysis (millions of plaintext-ciphertext pairs). Acquisition cost for SCA can be high but manageable. Imbalanced data in SCA, classes (key byte values) are naturally balanced. In distinguisher training, balancing "cipher" vs. "random" instances is crucial. Preventing overfitting techniques include dropout layers, L1/L2 regularization, early stopping, and rigorous validation on held-out datasets. Transfer learning potential to leverage models pre-trained on similar tasks or devices to reduce data needs, though exploration is limited.

4. DL FOR SIDE-CHANNEL ANALYSIS (DL-SCA)

4.1. The SCA landscape and DL's entry point

Traditional SCA template attacks (optimal profiled), CPA (popular non-profiled), DPA. Limitations Sensitivity to trace misalignment, noise, countermeasures; need for manual feature selection (Points-of-interest (POIs)); performance degradation with complex leakage functions. DL revolution Maghrebi *et al*. [31] and Cagli *et al*. [8] were among the first to demonstrate CNNs outperforming classical profiled SCA, handling raw traces directly.

4.2. Profiled DL-SCA the new gold standard

Workflow i) Acquire profiling traces (known inputs, known keys). ii) Train DL model (e.g., CNN) to predict intermediate values (e.g., S-box output) or directly key bytes from traces. iii) Acquire attack traces (known inputs, unknown key). iv) Use trained model to predict key candidates. v) Rank/recover key. Superior performance consistently shown to achieve lower NTD/GE than Template Attacks, especially in noisy environments or with misalignment. Kim *et al*. [9] explicitly demonstrated CNNs effectively utilize noise. Raw trace processing a major advantage is bypassing the need for explicit POI selection or trace alignment. CNNs learn robust features invariant to small shifts. Multi-task learning training a single model to predict multiple intermediate values or key bytes simultaneously, improving efficiency.

4.3. Combating countermeasures with DL

Jitter and random delays Cagli *et al*. [8] pioneered using data augmentation during training. By artificially adding random temporal shifts (jitter) to profiling traces, the CNN learns to be invariant to such countermeasures, outperforming classical attacks requiring explicit realignment. Desynchronization Similar to jitter, handled effectively by data augmentation (shifting) and CNN's inherent shift-invariance. Masking More challenging, DL-SCA can potentially learn higher-order moments or complex interactions between shares. Requires more traces and potentially more complex models or specific pre-processing. State-of-the-art but still an active challenge compared to unmasked scenarios. Hiding (Noise addition) CNNs demonstrate inherent robustness to Gaussian noise. Performance degrades gracefully compared to classical attacks.

4.4. Non-profiled and semi-profiled DL-SCA

Non-profiled challenges significantly harder than profiled. Requires unsupervised or self-supervised learning, dimensionality reduction (PCA, autoencoders), or clustering guided by DL features. Semi-profiled approaches utilizing some limited knowledge or model from a similar device. An area of ongoing research with potential for reducing profiling costs [13].

4.5. Systematization and comparative analysis

Surveys by Hettwer *et al*. [32], Masure *et al*. [33], and the comprehensive SoK by Picek *et al*. [13] provide invaluable taxonomies, classifying DL-SCA works by attack type (profiled/non-profiled), target (algorithm, implementation), DL architecture, countermeasures addressed, evaluation metrics. Highlighting the evolution from MLPs to CNNs as the dominant architecture. Emphasizing the critical importance of rigorous evaluation methodologies, reporting NTD/GE/SR curves, and using multiple datasets. Table 1 presents a structured taxonomy of DL-based SCA approaches, summarizing attack models, target algorithms, neural architectures, countermeasures addressed, and evaluation metrics reported in the literature.

Table 1. Detailed taxonomy of DL-SCA approaches (adapted from [13], [32], [33])

Characteristic	Categories	Examples/Notes	References
Attack type	Profiled, non-profiled, semi-profiled	Profiled dominates research and performance. Non-profiled remains challenging.	[8], [9], [13], [31]-[33]
Target algorithm	AES, DES, RSA, ECC, Lightweight (PRESENT, SIMON, SPECK), ASCON, etc.	AES is the most common benchmark. Lightweight ciphers are frequent targets due to constrained implementations.	[8], [9], [11]-[13], [32]
Leakage source	Power consumption (DPA), electromagnetic (EMA), timing, photonic, acoustic	Power and EM are most prevalent. DL shows promise for others.	[9], [13], [32], [33]
DL architecture	MLP, CNN (1D, 2D), ResNet, VGG, Autoencoders, RNN/LSTM (rare), hybrid models	CNN (1D) is state-of-the-art for profiled SCA on raw traces. MLP used earlier/hybrid. Autoencoders explored for non-profiled/feature reduction.	[8], [9], [11]-[13], [31], [33]
Countermeasures addressed	None, jitter/random delays, masking (1st order, higher-order), shuffling, hiding	Data Augmentation highly effective against jitter/delays/shuffling. Masking remains a significant challenge requiring more traces/complex models.	[8], [9], [12], [24], [26]
Key contribution focus	Raw trace processing, robustness to noise, countermeasure resilience, few-shot learning, model interpretability	[9] (Noise), [8] (Jitter), [11] (Algorithmic link), [13], [33] (Systematization), nascent work on explainable AI (XAI).	[8]-[10], [12], [26]
Evaluation metrics	SR, GE, NTD	Reporting full SR/GE/NTD curves is essential. NTD @ SR=80-90% common.	[8], [9], [13], [32], [33]

5. DL FOR CRYPTANALYSIS OF SYMMETRIC PRIMITIVES

5.1. Motivation and scope

While DL-SCA targets implementations, DL is also applied to attack the mathematical structure of symmetric ciphers (block ciphers, stream ciphers). Goals include distinguishers differentiating the cipher from a random permutation with fewer samples than brute-force. Improved differential/linear cryptanalysis finding better differential characteristics or linear approximations [34]. Key recovery directly or indirectly reducing the complexity of key search.

5.2. Neural distinguishers

Concept train a DL model (MLP, CNN) to classify tuples (e.g., pairs of plaintexts and corresponding ciphertexts, or input/output differences) as originating from the target cipher or a random permutation [7], [11], [12]. PRESENT case study Mishra *et al*. [12] demonstrated that MLPs could effectively distinguish reduced-round versions (e.g., 5-7 rounds) of the lightweight cipher PRESENT from random with high accuracy. This demonstrated DL's ability to capture non-randomness arising from the cipher's structure, albeit for rounds well below the full 31 rounds. Their work highlighted the potential and the significant challenge of scaling. Methodology requires generating vast datasets of cipher and random permutation outputs. Careful balancing is crucial. Evaluation involves accuracy on a separate test set and often calculating the advantage over random guessing.

5.3. DL-enhanced differential cryptanalysis Gohr's landmark

Attack Speck32/64 Gohr [11] presented a groundbreaking application of CNNs to differential cryptanalysis of the lightweight ARX cipher Speck32/64. Methodology i) Train a CNN to predict the output

difference of the last round given the input difference to the first round (or a differential characteristic over multiple rounds). ii) Use this neural distinguisher to filter ciphertext pairs more likely to follow a high-probability differential characteristic. iii) Use the filtered pairs in a classical key recovery attack on the last round(s). Results Achieved significant improvements over classical differential attacks on round-reduced Speck (e.g., up to 11 rounds). The CNN learned complex differential properties directly from data, effectively automating a crucial part of the cryptanalytic process. This demonstrated DL's potential to improve traditional cryptanalytic techniques. Significance bridged the gap between "academic" neural distinguishers and practical key recovery, albeit on reduced rounds of a lightweight cipher.

5.4. Critical examination of neural cryptanalysis

Benamira *et al*. [7] provide a crucial critical perspective distinguisher vs. key recovery emphasize that a good neural distinguisher does not automatically translate into an efficient key recovery attack. Gohr's work is notable for making this link explicit. The "why" question critically question *what* the neural network is actually learning. Is it capturing fundamental cryptographic properties (like high-probability differentials) or superficial patterns specific to the training data distribution? Lack of explainability is a major concern. Data efficiency and generalization neural distinguishers often require enormous datasets (millions of samples) and their generalization to different key classes or slightly modified ciphers is not guaranteed. Scalability the central challenge remains scaling these approaches to full-round versions of standardized, secure ciphers like AES-128. Current successes are largely confined to lightweight ciphers or severely reduced rounds. Singh *et al.* [6] also highlight this limitation. Table 2 highlights representative DL applications in the cryptanalysis of symmetric primitives and summarizes key findings and limitations.

Table 2. DL applications in symmetric primitive cryptanalysis - extended analysis

Target	Technique	Key finding	Limitations/Challenges	References
PRESENT (reduced)	MLP Distinguisher	Demonstrated capability of MLPs to distinguish 5-7 rounds of PRESENT from random with high accuracy. Validated DL's ability to detect non-randomness.	Accuracy drops significantly towards full (31) rounds. Limited to distinguisher, not key recovery. Lack of insight into *why*.	[12]
Speck32/64 (reduced)	CNN Differential Aid	CNN learned to predict differential characteristics, significantly improving key recovery on up to 11 rounds vs. classical differential cryptanalysis.	Attack complexity still high. Full cipher (22/23 rounds) remains secure. Requires massive training data (100M+ samples).	[11]
General distinguishers	MLP/CNN	Multiple studies confirm DL's ability to find statistical distinguishers for various reduced-round ciphers.	Practical impact on full ciphers minimal so far. High data requirements. Explainability gap ("black box").	[6], [7], [12]
Key recovery	Hybrid (DL + classical)	Most promising approach Use DL distinguisher to filter/rank candidates for classical key recovery steps (like Gohr [11]). Direct DL key prediction is infeasible for secure ciphers.	Integration complexity. Requires careful design. Still limited to reduced rounds/lightweight targets. Scaling is the core issue.	[7], [11]

6. CRITICAL ANALYSIS AND LIMITATIONS

6.1. The explainability ("Black Box") problem

This is arguably the most significant criticism and limitation [35], [36]. Lack of Insight DL models, especially complex CNNs, are opaque. It's extremely difficult to understand why a model makes a specific prediction or what specific cryptographic property it has learned [5], [7], [33]. This hinders gaining new fundamental cryptographic knowledge from successful attacks [37]. Trusting the model's output, especially in security-critical contexts. Debugging or improving models systematically. Addressing XAI in cryptanalysis applying XAI techniques (saliency maps, LIME, SHAP) to cryptanalysis models is nascent and challenging due to the complexity and sensitivity of cryptographic functions. Benamira *et al*. [7] strongly advocate for research into interpretable models or methods to extract human-understandable rules from DL distinguishers. DL models used in cryptanalysis-particularly CNNs-often behave as black boxes, making it difficult to explain why a prediction is made or which cryptographic/leakage features are exploited. This limits the extraction of human-interpretable rules, systematic debugging, and trust in security-critical settings. Accordingly, recent work advocates integrating XAI tools (e.g., saliency analysis and attribution methods) to identify influential samples/operations and to connect empirical success with cryptographic insight [7], [33], [35].

6.2. Scalability to secure primitives

Scalability remains the primary barrier for DL-assisted algorithmic cryptanalysis. Current successes mainly target reduced-round configurations and lightweight designs; extending neural distinguishers or DL-aided differentials to full-round standardized primitives faces an exponential growth in complexity and data requirements and remains an open challenge [6], [7], [38]. In DL-SCA, the analogous frontier is efficient key recovery under high-order masking, very low SNR, and hardened secure hardware. From a deployment perspective, practical SCA pipelines may also rely on edge/cloud offloading for trace processing and inference, where latency and service placement constraints become relevant [39].

7. FUTURE RESEARCH DIRECTIONS

7.1. XAI for cryptanalysis

Developing specialized XAI techniques to interpret what DL models learn about cryptographic algorithms or leakage functions. Designing inherently more interpretable DL architectures for cryptanalysis. Using insights from XAI to guide traditional cryptanalysis or improve cipher designs [40], [41]. Priority should be given to XAI methods tailored to cryptographic and leakage domains to reveal which operations, time samples, or bits drive model decisions and to extract human-understandable rules from trained models.

7.2. Bridging the theory-practice gap

A major open challenge is to translate empirical neural distinguishers into theory-backed cryptanalytic guarantees. Bridging this theory–practice gap requires clearer connections between what a model learns, the statistical advantage it achieves, and how this can be converted into practical key-recovery or proof of weakness [42].

7.3. Improving data efficiency and generalization

Transfer learning adapting models pre-trained on one device/cipher/task to perform well on a related target with minimal new data [43]. Few-shot/meta-learning enabling models to learn effectively from very few examples, crucial for attacking rare or highly secured devices [44]. Synthetic data generation exploring high-fidelity simulation or generative models (GANs, diffusion models) to create realistic training data, reducing reliance on physical acquisition, especially for algorithmic attacks [45]. Self-supervised learning Leveraging unlabeled data for pre-training, reducing the burden of labeled data acquisition. Transfer learning, few-shot/meta-learning, and high-fidelity synthetic data generation (e.g., GANs/diffusion) can reduce profiling cost and improve robustness across devices and noise regimes.

7.4. Advanced architectures and techniques

Graph neural networks, transformers, reinforcement learning, and ensemble/fusion approaches are promising for modeling structured dependencies in cipher states or long traces and for automating parts of the cryptanalytic search process [46]–[48]. Extending neural cryptanalysis beyond reduced-round block ciphers to additional primitives and leakage modalities is a promising direction but requires careful threat modeling and scalable training pipelines [49].

7.5. Standardization and reproducibility

Community benchmarks, standardized evaluation protocols, and reproducible open-source tooling are essential to enable fair comparison and accelerate progress. Complementary to empirical benchmarking, formal methods for verification/certification of implementations and defenses can strengthen confidence in security claims, especially when ML components are used in the attack or defense loop [50], [51].

8. CONCLUSION

DL has significantly reshaped the landscape of modern cryptanalysis. In the domain of SCA, DL-particularly CCNs-has emerged as the state of the art for profiled attacks, enabling direct processing of raw traces, improved robustness to noise and common countermeasures such as jitter, and effective key recovery with fewer traces. These advances, demonstrated in works such as [8], [9], and systematized in [6], [13], [32], represent a substantial practical improvement over classical techniques.

For algorithmic cryptanalysis of symmetric primitives, DL has shown promise in automating tasks such as statistical distinguishers and the discovery of differential characteristics. Notably, Gohr's work on Speck [11] illustrates how neural models can enhance classical cryptanalytic workflows and improve key recovery on reduced-round ciphers. However, current successes remain largely confined to lightweight

designs or reduced-round variants, and scaling such approaches to full-round standardized ciphers continues to be an open and challenging problem.

This review highlights limitations that constrain broader impact: the black-box nature of deep models limits explainability and the extraction of cryptographic insight, while data dependency and generalization issues hinder robustness across devices, implementations, and attack scenarios. Additionally, computational costs and the need for rigorous, reproducible evaluation frameworks remain important concerns. Future research should therefore focus on improving explainability through XAI, bridging the theory-practice gap, enhancing data efficiency and generalization, and exploring novel architectures and hybrid strategies capable of addressing more complex cryptographic targets. While DL is not a universal solution to cryptanalysis, its ability to automate complex pattern recognition ensures that it will remain a powerful and evolving tool within the ongoing arms race between cryptographic design and cryptanalytic attack.

FUNDING INFORMATION

The authors state no funding is involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Oussama Noui	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	
Amine Barkat		✓				✓		✓	✓	✓	✓	✓		

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

The authors state no conflict of interest.

REFERENCES

[1] F. E. Potestad-Ordonez, E. Tena-Sanchez, A. J. Acosta-Jimenez, C. J. Jimenez-Fernandez, and R. Chaves, "Design and evaluation of countermeasures against fault injection attacks and power side-channel leakage exploration for AES block cipher," *IEEE Access*, vol. 10, pp. 65548–65561, 2022, doi: 10.1109/ACCESS.2022.3183764.

[2] T. Sugawara, N. Shoji, K. Sakiyama, K. Matsuda, N. Miura, and M. Nagata, "Side-channel leakage from sensor-based countermeasures against fault injection attack," *Microelectronics Journal*, vol. 90, pp. 63–71, 2019, doi: 10.1016/j.mejo.2019.05.017.

[3] N. Benhadjyoussef, M. Karmani, and M. Machhout, "Power-based side channel analysis and fault injection: hacking techniques and combined countermeasure," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 5, pp. 709–719, 2021, doi: 10.14569/IJACSA.2021.0120583.

[4] S. Naserelden, N. Alias, A. Altigani, A. Mohamed, and S. Badreddine, "Advance attacks on AES: a comprehensive review of side channel, fault injection, machine learning and quantum techniques," *Edelweiss Applied Science and Technology*, vol. 9, no. 4, pp. 2471–2486, 2025, doi: 10.55214/25768484.v9i4.6586.

[5] N. Mohamed, "Artificial intelligence and machine learning in cybersecurity: a deep dive into state-of-the-art techniques and future paradigms," *Knowledge and Information Systems*, vol. 67, no. 8, pp. 6969–7055, 2025, doi: 10.1007/s10115-025-02429-y.

[6] A. Singh, K. B. Sivangi, and A. N. Tentu, "Machine learning and cryptanalysis: an in-depth exploration of current practices and future potential," *Journal of Computing Theories and Applications*, vol. 1, no. 3, pp. 257–272, 2024, doi: 10.62411/jcta.9851.

[7] A. Benamira, D. Gerault, T. Peyrin, and Q. Q. Tan, "A deeper look at machine learning-based cryptanalysis," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 12696 LNCS, pp. 805–835, 2021, doi: 10.1007/978-3-030-77870-5_28.

[8] E. Cagli, C. Dumas, and E. Prouff, "Convolutional neural networks with data augmentation against jitter-based countermeasures: Profiling attacks without pre-processing," *Lecture Notes in Computer Science*, vol. 10529 LNCS, pp. 45–68, 2017, doi: 10.1007/978-3-319-66787-4_3.

[9] 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, doi: 10.46586/tches.v2019.i3.148-179.

[10] Purwono, A. Ma'arif, W. Rahmaniar, H. I. K. Fathurrahman, A. Z. K. Frisky, and Q. M. U. Haq, "Understanding of Convolutional Neural Network (CNN): a review," *International Journal of Robotics and Control Systems*, vol. 2, no. 4, pp. 739–748, 2022, doi: 10.31763/ijrcs.v2i4.888.

[11] A. Gohr, "Improving attacks on round-reduced speck32/64 using deep learning," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11693 LNCS, pp. 150–179, 2019, doi: 10.1007/978-3-030-26951-7_6.
[12] G. Mishra, S. V. S. S. N. V. G. Krishna Murthy, and S. K. Pal, "Neural network based analysis of lightweight block cipher PRESENT," *Advances in Intelligent Systems and Computing*, vol. 741, pp. 969–978, 2019, doi: 10.1007/978-981-13-0761-4_91.
[13] S. Picek, G. Perin, L. Mariot, L. Wu, and L. Batina, "SoK: deep learning-based physical side-channel analysis," *ACM Computing Surveys*, vol. 55, no. 11, 2023, doi: 10.1145/3569577.
[14] R. A. de Oliveira and M. H. J. Bollen, "Deep learning for power quality," *Electric Power Systems Research*, vol. 214, 2023, doi: 10.1016/j.epsr.2022.108887.
[15] K. Sharifani and M. Amini, "Machine learning and deep learning: a review of methods and applications," *World Information Technology and Engineering Journal*, vol. 10, no. 7, pp. 3898–3904, 2023, [Online]. Available: https://ssrn.com/abstract=4458723.
[16] A. Mehrish, N. Majumder, R. Bharadwaj, R. Mihalcea, and S. Poria, "A review of deep learning techniques for speech processing," *Information Fusion*, vol. 99, 2023, doi: 10.1016/j.inffus.2023.101869.
[17] S. J. D. Prince, *Understanding deep learning*, no. May 29. Cambridge: MIT Press, 2025.
[18] D. Z. Chen, F. Trevizan, and S. Thiébaux, "Return to tradition: learning reliable heuristics with classical machine learning," *Proceedings International Conference on Automated Planning and Scheduling, ICAPS*, vol. 34, pp. 68–76, 2024, doi: 10.1609/icaps.v34i1.31462.
[19] E. Brier, C. Clavier, and F. Olivier, "Correlation power analysis with a leakage model," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 3156, pp. 16–29, 2004, doi: 10.1007/978-3-540-28632-5_2.
[20] P. Bottinelli and J. W. Bos, "Computational aspects of correlation power analysis," *Journal of Cryptographic Engineering*, vol. 7, no. 3, pp. 167–181, 2017, doi: 10.1007/s13389-016-0122-9.
[21] N. Oussama, B. Assia, and N. Lemnouar, "Secure image encryption scheme based on polar decomposition and chaotic map," *International Journal of Information and Communication Technology*, vol. 10, no. 4, pp. 437–453, 2017, doi: 10.1504/IJICT.2017.084339.
[22] N. Oussama and N. Lemnouar, "A robust watermarking scheme for ownership protection and deadlock prevention," *ACM International Conference Proceeding Series*, vol. 23-25-Nove, 2015, doi: 10.1145/2816839.2816915.
[23] O. Harkati, L. Noui, and A. Beloucif, "A lossless probabilistic image encryption algorithm based on QSD decomposition and matrix transformations," *International Journal of Computers and Applications*, vol. 46, no. 10, pp. 868–879, 2024, doi: 10.1080/1206212X.2024.2392168.
[24] T. D. Barve, A. Y. Samant, and M. S. Kulaye, "A comparative study of optimization algorithms in deep learning: SGD, Adam, And Beyond," *International Journal of Computer Technology and Electronics Communication (IJCTEC)*, vol. 9001, no. 4, p. 11018, 2025, [Online]. Available: www.ijctece.com.
[25] M. M. Taye, "Understanding of machine learning with deep learning: architectures, workflow, applications and future directions," *Computers*, vol. 12, no. 5, 2023, doi: 10.3390/computers12050091.
[26] R. Kruse, S. Mostaghim, C. Borgelt, C. Braune, and M. Steinbrecher, "Multi-layer perceptrons," pp. 53–124, 2022, doi: 10.1007/978-3-030-42227-1_5.
[27] I. El Jaafari, A. Ellahyani, and S. Charfi, "Rectified non-linear unit for convolution neural network," *Journal of Physics: Conference Series*, vol. 1743, no. 1, 2021, doi: 10.1088/1742-6596/1743/1/012014.
[28] A. J. Cabrera-Gutierrez, E. Castillo, A. Escobar-Molero, J. A. Alvarez-Bermejo, D. P. Morales, and L. Parrilla, "Integration of hardware security modules and permissioned blockchain in industrial IoT networks," *IEEE Access*, vol. 10, pp. 114331–114345, 2022, doi: 10.1109/ACCESS.2022.3217815.
[29] S. H. Wang, M. A. Khan, and Y. D. Zhang, "VISPNN: VGG-inspired stochastic pooling neural network," *Computers, Materials and Continua*, vol. 70, no. 2, pp. 3081–3097, 2022, doi: 10.32604/cmc.2022.019447.
[30] J. A Ilemobayo *et al.*, "Hyperparameter tuning in machine learning: a comprehensive review," *Journal of Engineering Research and Reports*, vol. 26, no. 6, pp. 388–395, 2024, doi: 10.9734/jerr/2024/v26i61188.
[31] H. Maghrebi, T. Portigliatti, and E. Prouff, "Breaking cryptographic implementations using deep learning techniques," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10076 LNCS, pp. 3–26, 2016, doi: 10.1007/978-3-319-49445-6_1.
[32] B. Hettwer, S. Gehrer, and T. Güneysu, "Applications of machine learning techniques in side-channel attacks: a survey," *Journal of Cryptographic Engineering*, vol. 10, no. 2, pp. 135–162, 2020, doi: 10.1007/s13389-019-00212-8.
[33] L. Masure, C. Dumas, and E. Prouff, "A comprehensive study of deep learning for side-channel analysis," *IACR Transactions on Cryptographic Hardware and Embedded Systems*, vol. 2020, no. 1, pp. 348–375, 2020, doi: 10.13154/tches.v2020.i1.348-375.
[34] R. Watanabe, N. Ghafoori, and A. Miyaji, "Improved differential-linear cryptanalysis of reduced rounds of ChaCha," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 14402 LNCS, pp. 269–281, 2024, doi: 10.1007/978-981-99-8024-6_21.
[35] Y. T. Goi, S. M. Leong, R. C. W. Phan, S. Lai, and A. Salagean, "Unveiling the black box: neural cryptanalysis with XAI," *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*, pp. 1951–1956, 2024, doi: 10.1109/SMC54092.2024.10831564.
[36] Z. Tolba, M. Derdour, M. A. Ferrag, S. M. Muyeen, and M. Benbouzid, "Automated deep learning BLACK-BOX attack for multimedia P-BOX SECURITY ASSESSMENT," *IEEE Access*, vol. 10, pp. 94019–94039, 2022, doi: 10.1109/ACCESS.2022.3204175.
[37] N. Lemnouar, "Security limitations of Shamir's secret sharing," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 26, no. 4, pp. 977–989, 2023, doi: 10.1080/09720529.2021.1961902.
[38] K. Jang *et al.*, "Quantum implementation and analysis of SHA-2 and SHA-3," *IEEE Transactions on Emerging Topics in Computing*, vol. 13, no. 3, pp. 919–934, 2025, doi: 10.1109/TETC.2025.3546648.
[39] K. Gasmi, S. Dilek, S. Tosun, and S. Ozdemir, "A survey on computation offloading and service placement in fog computing-based IoT," *Journal of Supercomputing*, vol. 78, no. 2, pp. 1983–2014, 2022, doi: 10.1007/s11227-021-03941-y.
[40] T. Senevirathna, V. H. La, S. Marcha, B. Siniarski, M. Liyanage, and S. Wang, "A survey on XAI for 5G and Beyond security: technical aspects, challenges and research directions," *IEEE Communications Surveys and Tutorials*, vol. 27, no. 2, pp. 941–973, 2025, doi: 10.1109/COMST.2024.3437248.

[41] H. Manthena, S. Shajarian, J. C. Kimmell, M. Abdelsalam, S. Khorsandroo, and M. Gupta, "Explainable artificial intelligence (XAI) for malware analysis: a survey of techniques, applications, and open challenges," *IEEE Access*, vol. 13, pp. 61611–61640, 2025, doi: 10.1109/ACCESS.2025.3555926.
[42] R. Law, S. S. I. Lei, K. Zhang, and A. Lau, "Bridging the theory-practice gap: a critical reflection on information and communication technology research," *International Journal of Contemporary Hospitality Management*, vol. 36, no. 6, pp. 1980–1990, 2024, doi: 10.1108/IJCHM-02-2023-0131.
[43] J. Hao, "Deep learning-based medical image analysis with explainable transfer learning," *Proceedings - 2023 International Conference on Computer Engineering and Distance Learning, CEDL 2023*, pp. 106–109, 2023, doi: 10.1109/CEDL60560.2023.00029.
[44] L. Qiao, Y. Zhang, Q. Wang, D. Li, and S. Peng, "Fault diagnosis for wind turbine generators based on model-agnostic meta-learning: a few-shot learning method," *Expert Systems with Applications*, vol. 267, 2025, doi: 10.1016/j.eswa.2024.126171.
[45] S. Bengesi, H. El-Sayed, M. K. Sarker, Y. Houkpati, J. Irungu, and T. Oladunni, "Advancements in generative AI: a comprehensive review of GANs, GPT, autoencoders, diffusion model, and transformers," *IEEE Access*, vol. 12, pp. 69812–69837, 2024, doi: 10.1109/ACCESS.2024.3397775.
[46] P. Kisanga, I. Woungang, I. Traore, and G. H. S. Carvalho, "Network anomaly detection using a graph neural network," *2023 International Conference on Computing, Networking and Communications, ICNC 2023*, pp. 61–65, 2023, doi: 10.1109/ICNC57223.2023.10074111.
[47] D. Pecioski, V. Gavriloski, S. Domazetovska, and A. Ignjatovska, "An overview of reinforcement learning techniques," *12th Mediterranean Conference on Embedded Computing, MECO 2023*, 2023, doi: 10.1109/MECO58584.2023.10155066.
[48] W. Li, Y. Peng, M. Zhang, L. Ding, H. Hu, and L. Shen, "Deep model fusion: a survey," *IEEE Transactions on Neural Networks and Learning Systems*, 2025, doi: 10.1109/TNNLS.2025.3628666.
[49] A. V. Shoshina, G. I. Borzunov, and E. Y. Ivanova, "Application of bio-inspired algorithms to the cryptanalysis of asymmetric ciphers on the basis of composite number," *Proceedings of the 2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering, ElConRus 2021*, pp. 2399–2403, 2021, doi: 10.1109/ElConRus51938.2021.9396242.
[50] A. Swaroop *et al.*, "A comprehensive overview of formal methods and deep learning for verification and optimization," *2024 International Conference on Decision Aid Sciences and Applications, DASA 2024*, 2024, doi: 10.1109/DASA63652.2024.10836654.
[51] C. Urban and A. Miné, "A review of formal methods applied to machine learning," 2021, [Online]. Available: http://arxiv.org/abs/2104.02466.

BIOGRAPHIES OF AUTHORS

Oussama Noui is an associate professor at the University of Batna 1, Algeria. He holds a Ph.D. degree in Computer Science with specialization in Cryptography and Security. His research areas are authentication, image encryption, and watermarking. He can be contacted at email: oussama.noui@univ-batna.dz.

Amine Barkat has a Ph.D. from Politecnico Di Milano with a main focus on Cloud Computing systems, virtualization, and mobile networks (4G and 5G). Currently, he is an associate professor at the Higher National School of Renewable Energies, Environment and Sustainable Development, Algeria. With a long background in various technologies and experience with universities and companies in more than 15 European countries, Amine delivers training and professional services for cloud-related topics, including data center virtualization, NFV, and storage networks. The delivered services include different business levels, mainly pre-sales and post-sales. He can be contacted at email: amine.barkat@polimi.it.