Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 4, 2471-2486
2025
Publisher: Learning Gate
DOI: 10.55214/25768484.v9i4.6586

Advance attacks on AES: A comprehensive review of side channel, fault injection, machine learning and quantum techniques

Shiraz Naserelden[1]*, Norma Alias[1,2], Abdelrahamn Altigani[3], Ahmed Mohamed[3], Said Badreddine[3,4]
[1]Department of Mathematical Sciences, Faculty of Science, Universiti Teknologi Malaysia (UTM), Johor Bahru, 81310, Malaysia; heshiraz@live.utm.my (S.N.).
[2]Center of Engineering Education, Universiti Teknologi Malaysia (UTM), Johor Bahru, 81310, Malaysia.
[3]Computer Information Science, Higher Colleges of Technology, 25026, Al Ain, Abu Dhabi, UAE.
[4]School of Computing, Engineering & Digital Technologies, Teesside University, Middlesbrough TS1 3BX, UK.

Abstract: The Advanced Encryption Standard (AES) remains a foundational component of modern cryptography, securing vast volumes of digital communication and data storage. Despite its robust design and widespread adoption, AES continues to be the subject of intensive cryptanalytic research. This paper presents a review of recent advances in attacks against AES, categorizing them into four domains: side-channel attacks, fault injection attacks, machine learning and AI-based attacks, and quantum computing threats. For each category, representative studies published between 2021 and 2025 are analyzed with respect to methodology, data requirements, attack complexity, and practical applicability. The review highlights both vulnerabilities exposed in specific AES implementations and the evolution of attack methodologies, thereby providing a comprehensive perspective on the contemporary threat landscape. The findings underscore the need for continuous evaluation and adaptation of AES-based systems to ensure cryptographic resilience in the face of advancing adversarial capabilities.

Keywords: *Advanced encryption standard (AES), AES security, Side-channel attacks, Fault injection, Machine learning in cryptanalysis, Quantum cryptanalysis*

1. Introduction

The Advanced Encryption Standard (AES), establish by the National Institute of Standards and Technology (NIST) in 2001 [1, 2] has become the standard for symmetric-key encryption across a broad spectrum of applications, from secure internet transactions to classified government communications. Its architecture, based on a substitution-permutation network and a fixed number of rounds dependent on key length [3, 4] was designed to balance security with performance efficiency. While AES has withstood extensive cryptanalytic scrutiny over the past two decades, the rapid advancement of computational capabilities and the emergence of new cryptographic attack surfaces necessitate continuous revaluation of its security assurances [5].

Recent technological developments particularly in artificial intelligence, quantum computing, and the proliferation of embedded and Internet of Things (IoT) devices have significantly altered the threat landscape [6]. These advances have expanded the avenues available to adversaries, enabling novel classes of attacks that can exploit subtle implementation flaws, statistical leakages, or theoretical weaknesses. Between 2021 and 2025, an increasing number of studies have demonstrated the feasibility of attacks on AES implementations using deep learning, fault injection, and quantum algorithms, often targeting specific use cases in embedded systems or constrained hardware environments [7, 8].

History: Received: 24 February 2025; Revised: 8 April 2025; Accepted: 11 April 2025; Published: 26 April 2025
* Correspondence: heshiraz@live.utm.my

Given the growing complexity and scope of cryptanalytic research, there is a clear need for a comprehensive review that synthesizes recent findings and organizes them by methodological category. While individual papers have provided detailed insights into attack vectors, no unified study to date has thoroughly categorized and analyzed the post-2020 body of literature on AES cryptanalysis. This paper aims to address that gap by conducting a structured review of contemporary attacks on AES, focusing on side-channel attacks, fault injection strategies, machine learning-augmented approaches and deep learning attacks. Additionally, we examine the impact of emerging quantum techniques and explore the effectiveness of countermeasures currently deployed in practice.

By organizing recent work into clearly defined categories, this review seeks to provide cryptographers, security engineers, and researchers with an accessible reference to current threats and defences pertaining to AES. Furthermore, it aims to identify open challenges and research opportunities in the ongoing effort to secure symmetric-key cryptography in the post-quantum era.

2. Methodology

This review adopts a structured methodology consistent with best practices in literature reviews, with the aim of synthesizing recent cryptanalytic advances targeting the Advanced Encryption Standard (AES). The approach comprises three main stages: literature acquisition, classification of attack types, and synthesis of findings.

The objectives of this review are threefold:

1. To classify and summarize cryptanalytic techniques employed against AES in recent literature (2021–2025).
2. To analyse emerging patterns, trends, and methodological innovations in the attacks.
3. To provide insights into gaps in current research and suggest future directions for defence mechanisms and secure AES implementations.

To ensure a comprehensive and representative survey of relevant literature, we considered peer-reviewed journal articles, conference proceedings, and preprints from leading databases and repositories. These include the IACR Cryptology ePrint Archive, IEEE Xplore, SpringerLink, ScienceDirect, and Google Scholar. The search was conducted using a combination of keywords and Boolean logic to capture a wide range of AES-focused cryptanalysis research.

The search terms included but were not limited to:

- AES cryptanalysis
- Advanced Encryption Standard vulnerabilities
- side-channel attacks on AES
- fault injection AES
- quantum cryptanalysis AES
- machine learning AES attack
- deep learning side-channel AES

Each identified paper was reviewed for relevance, novelty, and technical depth. Eligible papers were then categorized based on their primary attack methodology. The four main categories are order to reflect both historical development and increasing attack complexity. We begin with Side-Channel Attacks bridge theory attacks like mathematical attacks to practice by exploiting physical implementations through passive observation. Fault Injection Attacks extend this physical paradigm to active manipulation representing a more invasive threat class. Emerging Machine Learning, Deep Learning & AI-Based Attacks follow, demonstrating how modern techniques enhance traditional methods (e.g., automating side-channel analysis or creating neural distinguishers). Finally, Quantum Computing Threats are discussed separately as forward-looking theoretical risks (e.g., Grover's algorithm), concluding with future challenges. This order mirrors the evolution of cryptographic attacks from pure mathematics to real-world exploitation, then to AI-augmented and quantum era threats while maintaining a pedagogical flow from basic to advanced concepts.

Hybrid approaches, such as Deep Learning-based cryptanalysis incorporating side-channel leakage or AI-enhanced fault analysis, were cross-referenced in multiple categories when applicable.

For each selected paper, we extracted relevant details including attack vector, complexity, required data (e.g., number of traces, faults, or oracle queries), target environment (software, hardware, or embedded), and the effectiveness or success metrics reported. This structured extraction supports the comparative analysis presented in the subsequent sections and facilitates an informed discussion of cryptanalytic trends and defence strategies.

3. Categories of Recent Attacks on AES

3.1. Side-Channel Attacks

Side-channel attacks (SCAs) constitute one of the most potent practical threats against AES implementations. Rather than targeting weaknesses in the cipher's mathematical structure, SCAs exploit physical leakage from the environment in which AES is executed such as variations in power consumption, amplitude-modulated electromagnetic (EM) emissions, execution timing, or cache behaviour. These leakages can inadvertently reveal sensitive intermediate values during encryption or decryption, ultimately compromising secret key material.

Recent advances between 2021 and 2025 have focused predominantly on power and EM-based SCAs, with a significant shift toward deep learning (DL)-based analysis. Traditional template attacks and correlation power analysis (CPA) have proven effective in many settings, but they often require extensive profiling and precise alignment of traces. In contrast, deep learning models have demonstrated robustness against misalignments and noise, while achieving high key recovery success rates even against masked implementations.

Kuroda, et al. [9] explores a non-profiled deep learning-based side-channel attack (DL-SCA) framework targeting masked AES-128 implementations, including Random Switching Masking (RSM). Their findings revealed that convolutional neural networks (CNNs) could still exploit subtle statistical dependencies in the masked traces to recover keys with high accuracy, especially when trained with sufficient data (e.g., 50K+ traces). Similarly, Fukuda, et al. [10] introduced a non-profiled deep learning side-channel attack technique multi-bit differential deep learning analysis (DDLA) targeting hardware-implemented AES with and without side-channel countermeasures that labels multiple intermediate bits simultaneously to increase learning efficiency. This multi-bit strategy proved effective against hardware implementations, achieving key recovery under realistic noise conditions.

Ghandali, et al. [11] proposed a profiled power side-channel attack using kernelized twin support vector machines (K-TSVM), which offered competitive performance to DL models with reduced training overhead. Other studies, such as those by Daehyeon Bae, et al. [12] and Negabi, et al. [13] focused on DL-powered power analysis attacks against masked AES on microcontrollers, notably the ATmega328P, highlighting the practical feasibility of DL-based SCA even on low-cost hardware platforms.

In parallel, EM-based SCAs have gained renewed interest. Wang [14] demonstrated that amplitude-modulated EM emissions from a provably masked AES implementation still exhibited exploitable leakage patterns. By using specialized probes and applying signal amplification techniques, key bytes were successfully extracted without violating the masking assumptions, illustrating that EM leakage can circumvent software-level protections.

Liu, et al. [15] introduced a novel frequency throttling side-channel attack, which leverages frequency scaling behaviour in CPUs to derive timing and power patterns associated with AES key-dependent execution paths. This attack poses a particular risk to embedded AES implementations on mobile or power-aware processors where dynamic frequency scaling is used.

Overall, side-channel attacks have evolved significantly due to machine learning integration. Despite countermeasures such as masking, shuffling, and hiding, modern SCA techniques especially DL-based approaches can learn complex leakage models that challenge traditional notions of leakage-resilience. These findings emphasize the need for stronger side-channel countermeasures, including hardware-level isolation, randomized S-box implementations, and higher-order masking schemes.

3.2. Fault Injection Attacks

Fault injection attacks (FIAs) are a class of physical attacks in which transient or permanent faults are deliberately introduced into cryptographic hardware or software implementations to bypass security mechanisms or extract sensitive information, such as secret keys. These attacks exploit the discrepancy between correct and faulty ciphertexts or internal computation states to infer critical information about the encryption process. Fault injection has proven to be a particularly effective attack vector against embedded systems and lightweight cryptographic implementations used in resource-constrained environments.

Recent research has proposed various techniques and countermeasures to both exploit and mitigate fault injection vulnerabilities in AES. Sheikhpour, et al. [16] designed a low-cost fault-resilient AES architecture (LC-FRAES) tailored for IoT environments. Their design shares hardware resources between encryption and decryption blocks and integrates strong error detection schemes, including precomputation with dynamically permuted operands (RDP) and temporal redundancy. LC-FRAES was extensively validated under diverse fault injection scenarios and demonstrated a high level of resilience against both transient and permanent faults. Moreover, the architecture achieved notable improvements in area efficiency, power consumption, and throughput, rendering it a practical and secure solution for real-world IoT deployments.

Pan, et al. [17] focused on enhancing the security of white-box AES implementations, which are particularly vulnerable to side-channel and fault attacks due to their exposure in software form. Their work combined a Differential Correlation Analysis (DCA) strategy with an intelligent fault injection algorithm, optimized using a genetic algorithm. This intelligent approach automated the parameter search for fault injection, improving attack efficiency by 980% compared to random selection. To counteract these attacks, the authors proposed a rotating S-box masking scheme, which proved effective in mitigating side-channel leakage and fault-induced weaknesses, thereby strengthening the white-box AES against combined DCA-FIA strategies.

Ghosal, et al. [18] proposed an enhancement to the AES MixColumns transformation MixColumn-Plus as a defence against Differential Fault Analysis (DFA) attacks. The method introduces two alternate transformation matrices, Mmix1 and Mmix2, which are incorporated into the MixColumns operation to increase the complexity of successful DFA attacks. The improved construction raises the attack complexity to approximately 2116, significantly hindering fault-based key recovery. Importantly, their FPGA-based hardware implementation showed that the security benefits were achieved with negligible impact on execution time, frequency, or hardware resource utilization, making it suitable for constrained embedded environments.

Further insights into DFA were provided by Anand, et al. [19] who demonstrated that fault injection into AES-based authenticated encryption schemes such as Rocca, Rocca-S, and AEGIS can drastically reduce the key search space. Their work underlines the necessity of robust fault detection and mitigation in schemes that rely on AES as a cryptographic core, especially in contexts requiring high security assurances.

Mestiri, et al. [20] introduced a parity-based fault detection scheme designed to enhance AES resilience in embedded systems. Their method computes and compares predicted and observed parities at each round of AES execution, incorporating hybrid redundancy mechanisms across multiple AES subcomponents, including the key scheduler, control logic, and round transformations. The design was implemented on a Xilinx Virtex-5 FPGA and demonstrated fault coverage of 99.999%, with moderate area overhead (45.59%) and minor performance degradation (3.83%). These results indicate a strong trade-off between cost and security, supporting the deployment of parity-based detection for critical embedded applications.

In summary, fault injection remains a potent threat to AES implementations, particularly in environments where physical access is feasible, or software-level protections are insufficient. However, recent advances in fault-aware hardware design, intelligent masking, and redundancy-based defenses have demonstrated considerable promise in mitigating these attacks. Moving forward, integrating these countermeasures into AES implementations especially those in IoT and mobile applications will be essential to maintaining robust fault tolerance and long-term cryptographic integrity.

3.3. Machine Learning and Artificial Intelligence based Attacks

The integration of artificial intelligence (AI) and machine learning (ML) techniques has significantly advanced the capabilities of cryptanalytic attacks against AES. These technologies have been employed both to automate and optimize traditional attack strategies and to uncover new vulnerabilities previously considered resistant to practical exploitation. While AI and ML are also being explored to enhance AES security through adaptive defence mechanisms, recent research predominantly emphasizes their offensive potential especially in the context of side-channel leakage, key guessing, and implementation-specific vulnerabilities.

In Zhong, et al. [21] introduced a novel key-guessing attack paradigm against logic block encryption, leveraging supervised ML techniques. Their methodology involved constructing a three-layer neural network coupled with a Naive Bayes classifier to approximate the inverse function of the encryption logic. The model was trained on data derived from simulations using encrypted netlists and activated integrated circuits (ICs), wherein randomized input and key vectors were applied to generate output responses. Once trained, the network effectively inferred key candidates, which were subsequently refined using the Naive Bayes classifier to improve prediction accuracy. Experimental results demonstrated that this approach could significantly accelerate key recovery, exposing practical weaknesses in contemporary logic encryption schemes.

Pasquale Arpaia [22] examined the effectiveness of machine learning-assisted side-channel attacks using power analysis. The study employed a multilayer perceptron (MLP) to classify power traces and recover the AES-128 secret key. Signal-to-noise ratio (SNR) analysis was used to identify points of interest (POIs) in the traces, which were then aligned with intermediate values derived from known plaintexts and masked keys. The model's performance was assessed using guessing entropy (GE) and its associated uncertainty, estimated via Monte Carlo simulations. Their experimental framework, validated on the ASCAD public dataset, revealed that ML can offer reliable key recovery performance even in the presence of side-channel countermeasures.

A different attack vector was demonstrated by Priya and Kapilamithran [23] who explored the intersection of generative AI and cryptographic implementation flaws, specifically targeting AES padding oracle vulnerabilities in web authentication systems. A custom login interface was developed using HTML, CSS, and JavaScript (frontend), with Python and Flask powering the backend. AES encryption was applied to user credentials before transmission. The authors exploited padding oracle vulnerabilities to infer plaintext lengths from ciphertexts and employed generative AI to predict potential password structures based on observed webpage behaviour. This combination of padding oracle exploitation and AI-enhanced input prediction illustrates a novel avenue for compromising AES-protected web applications.

In a broader context, Ahn, et al. [24] conducted a comprehensive survey examining vulnerabilities in widely used cryptographic network protocols such as TLS and SSH. While not limited to AES, the study analysed weaknesses across encryption schemes at the transport and application layers of the TCP/IP stack. Drawing from CVE reports and major scientific repositories, the survey categorized cryptographic vulnerabilities into groups targeting key exchange protocols, block and stream ciphers, and session establishment mechanisms. Their findings underscore the increasing complexity of attack surfaces in modern networked environments and provide recommendations for future research directions to mitigate cryptographic risk.

Despite their demonstrated success, AI and ML-based attacks on AES are not without significant limitations. First, training deep models such as MLPs and convolutional neural networks (CNNs) demands substantial computational resources, making them impractical for resource-constrained attackers or real-time applications. The efficacy of these models is highly contingent on the quality and volume of training data, with noise, desynchronization, and jitter in power traces reducing model accuracy. Identifying relevant POIs and extracting robust features for training often requires expert domain knowledge and time-consuming pre-processing.

Moreover, AES implementations frequently incorporate countermeasures such as masking, hiding, and randomization, which obfuscate side-channel leakage and degrade the performance of ML-based attacks. While machine learning can estimate the guessing entropy for partial key recovery, combining these estimates to recover the full 128-bit AES key remains computationally demanding. Many ML-based attacks focus on recovering subkeys (e.g., individual bytes), but the combinatorial complexity of reconstructing the complete key presents a persistent challenge.

Finally, the generalization and reproducibility of ML attacks across different devices and datasets remain open problems. Models trained on one platform may not transfer well to another due to variations in leakage patterns, hardware characteristics, and implementation-specific features. These factors highlight the need for ongoing research to improve the robustness, scalability, and transferability of ML-based attacks, as well as the development of resilient AES implementations capable of withstanding AI-assisted adversaries

3.3.1. Deep Learning Attacks

Deep learning (DL) has emerged as a transformative tool in the realm of cryptanalysis, offering enhanced capabilities for side-channel attacks on symmetric-key algorithms such as the Advanced Encryption Standard (AES). Leveraging the pattern recognition strength of neural networks, DL models can learn to extract cryptographic secrets from physical leakage data particularly power traces and electromagnetic emissions—thereby posing serious threats to the confidentiality of AES-based systems.

Among the most applied architectures are multi-layer perceptron's (MLPs) and convolutional neural networks (CNNs), which have demonstrated strong performance in both profiling and non-profiling side-channel attack scenarios. Profiling attacks involve training the model on a known device under controlled conditions to predict key-dependent leakages in a target system, while non-profiling attacks bypass the training phase on a reference device, instead attempting to infer patterns directly from the attack target. Numerous studies Daehyeon Bae, et al. [12]; Ghandali, et al. [11] and Swaminathan, et al. [25] have confirmed that DL-based approaches can achieve high accuracy in key recovery, even in the presence of countermeasures such as first-order masking.

Kubota, et al. [26] highlighted how CNNs could outperform classical side-channel techniques by automatically identifying points of interest (POIs) in traces, enabling the recovery of AES subkeys with minimal manual feature engineering. These models are particularly effective at learning complex, non-linear relationships between side-channel leakage and intermediate values within the AES round transformations.

Negabi, et al. [13] demonstrated a practical deep learning-based power analysis attack (PAA) against AES-128 implemented on an ATmega328P microcontroller. Their methodology involved training a CNN on captured power consumption traces from the target device, enabling the model to extract key-dependent features and ultimately reconstruct the AES-128 secret key. Notably, their approach achieved successful key recovery with fewer traces than traditional techniques, underscoring the efficiency and scalability of DL-based power analysis.

In another significant contribution, Tom´aˇs Gerlich [27] introduced DL-SITM a deep learning-augmented version of the See-in-the-Middle (SITM) attack. The authors generated plaintext pairs with controlled Hamming distances and analysed their differential power traces to localize collision points within AES's internal state. A CNN classifier trained on these differentials effectively automated what was previously a manual visual inspection step, enabling successful collision detection even under high-noise and jitter countermeasure conditions. DL-SITM demonstrated high precision and recall across multiple datasets, highlighting its robustness under realistic deployment settings.

Expanding the application scope of deep learning in cryptographic contexts, Dhanalakshmi [28] proposed an integrated DL-enhanced classification framework for wireless sensor networks (WSNs). The system utilized hybrid encryption combining AES and elliptic curve cryptography (HAE-ECC) to secure node communication. A multi-stage optimization and classification pipeline, incorporating Deep CNNs, Bi-directional LSTM, and an Attention Mechanism, was employed to identify attack behaviours in encrypted traffic. Although this approach focused more on intrusion detection than key recovery, it exemplifies how DL can complement cryptographic primitives to enhance overall system security.

Despite the promise shown by DL in cryptanalysis, several limitations constrain its practical applicability in real-world scenarios:

- Dependence on High-Quality Side-Channel Data: DL models require large, well-labelled datasets of power traces or EM emissions. Noise, desynchronization, and environmental variability can severely impact model performance.
- Significant Data and Computational Overhead: Training deep neural networks especially CNNs demands extensive computational resources and may not be feasible for real-time or constrained environments.
- Reliance on Known Plaintext or Ciphertext: Many DL-based attacks require access to input-output pairs, which may not always be available in practice.
- Reduced Effectiveness Against Countermeasures: Techniques such as masking, hiding, shuffling, and jitter injection can degrade the information content of traces, thereby lowering attack accuracy.
- Lack of Real-World Deployment Examples: While laboratory results are promising, few DL-based attacks have been documented in practical attack scenarios outside controlled experimental conditions.

These challenges highlight the dual necessity of advancing DL-based attack methodologies while simultaneously strengthening implementation-level countermeasures. Future work must focus on enhancing the generalizability, efficiency, and stealth of DL attacks, as well as building more resilient AES implementations across software and hardware platforms.

3.4. Quantum Computing Threats

The advent of quantum computing presents a significant paradigm shift in the field of cryptanalysis, with profound implications for both symmetric and asymmetric cryptographic systems. Although symmetric algorithms like AES are considered more resistant to quantum attacks than public-key schemes, they are by no means immune. Grover's algorithm enables a quadratic speedup in brute-force key search, effectively reducing the security level of AES-128 from 128 to 64 bits in a quantum context. This section surveys recent advances in quantum-assisted cryptanalysis techniques that target AES or AES-like primitives, focusing on structural, algebraic, and amplitude amplification-based methods.

Xiaoyang Dong [29] introduced an automated tool designed to optimize rebound attacks by identifying high-probability related-key differentials in AES-like constructions. The tool streamlines the process of modelling key differences and managing linear incompatibilities, which are prevalent in modern lightweight block ciphers. Applied to Saturnin (a NIST LWC Round 2 candidate), SKINNY-128-384, and Whirlpool, the framework yields 7- to 9-round collision attacks that exploit quantum collision finding principles without requiring qRAM or large memory. While not always faster than classical memory-intensive attacks, these techniques demonstrate practical relevance in quantum-ready cryptanalytic frameworks.

Chang, et al. [30] explored the vulnerability of AES-OTR (Offset Two-Round), an authenticated encryption algorithm, in the presence of quantum adversaries. By leveraging the Simon algorithm, a quantum algorithm suitable for finding hidden periodicities, the authors constructed a collision forgery attack capable of identifying non-trivial periods in associated data. The attack operates with near-certain success probability, underscoring the fragility of mode-specific AES constructions when evaluated under quantum assumptions. These findings suggest that even secure AES instantiations may require redesign in the post-quantum landscape.

Jang, et al. [31] reviewed the application of quantum meet-in-the-middle attacks, particularly quantum-enhanced adaptations of the classical Demirci–Selçuk framework. Although these attacks do not currently outperform Grover-based brute-force search in the case of full-round AES, they highlight structural vulnerabilities that could become exploitable in reduced-round or misconfigured implementations. Importantly, these techniques provide insight into how classical cryptanalytic frameworks may be augmented with quantum speedups to target specific components of AES.

Cai, et al. [32] proposed a quantum slide attack on 1K-AES using Quantum Amplitude Amplification (QAA). The method begins by generating a set of random plaintext-ciphertext pairs and partially decrypting them to derive XOR differentials. A QAA-based search is then employed to identify matching entries from a precomputed transformation table. Unlike Grover's algorithm, QAA is optimized for non-uniform superpositions and can efficiently handle scenarios with an unknown number of valid states. This makes it particularly effective in scenarios where the probability distribution of successful matches is sparse or non-deterministic.

Mandal, et al. [33] presented an implementation of Grover's algorithm using an enhanced Variational Quantum Amplitude Amplification (VQAA) technique, applied to symmetric key cryptanalysis. Their hybrid quantum-classical architecture utilizes a variational quantum circuit (VQC) to generate candidate key superpositions, with a cost function comparing trial encryptions to known ciphertexts. The optimization is guided by classical methods such as conjugate gradient descent. Their experiments, conducted in a simulated noise-free environment using Qiskit on a MacBook Pro (Apple M1 Pro), demonstrated significant qubit and depth reductions compared to standard Grover circuits. While the study focused on simplified ciphers like S-AES and Blowfish, the techniques are generalizable to larger-key AES variants.

The cumulative implications of Grover's algorithm for symmetric encryption are well-documented [33-35]. Specifically, Grover reduces the complexity of exhaustive key search from O(2n) to O(2n/2) where n is the key length. For AES-128, this implies an effective security strength of only 64 bits against a quantum-capable adversary. Consequently, NIST and other standards bodies now recommend the use of AES-256 for long-term post-quantum applications. While no quantum algorithm is currently known to break AES completely in polynomial time, the reduction in brute-force resistance necessitates a revaluation of key sizes and reinforces the need for quantum-resilient implementations.

In conclusion, while AES remains secure under current quantum threat models, emerging research illustrates how quantum techniques can systematically undermine various structural and operational aspects of AES and its variants. Continued development of quantum-resistant symmetric algorithms, along with hybrid defence strategies, will be essential to ensuring cryptographic robustness in the forthcoming post-quantum era.

4. Discussion and Limitations

The analysis of 25 recent studies on AES cryptanalysis reveals a clear evolution in attack strategies, reflecting both advances in attacker capabilities and the growing complexity of AES implementations across different platforms. These attacks span a wide spectrum from side-channel and fault injection to machine learning and quantum cryptanalysis each exhibiting distinct requirements, assumptions, and levels of practicality.

Side-channel attacks remain among the most prevalent and practically demonstrable threats. The review shows that modern DL-based side-channel attacks (e.g., CNNs, K-TSVMs) require tens of thousands of traces and are often effective against both masked and unmasked AES implementations. However, their success depends heavily on data quality, proper alignment, and model tuning. EM-based side-channel attacks similarly expose vulnerabilities in hardware-level AES even under masking, though they require careful probe placement and noise handling.

Fault injection attacks have matured considerably, as evidenced by their ability to bypass or degrade the security of embedded AES through glitching, electromagnetic interference, or induced hardware faults. The reviewed works demonstrate the effectiveness of intelligent fault injection algorithms using evolutionary techniques (e.g., genetic algorithms) and highlight countermeasures like MixColumn-Plus and parity-based detection as viable hardware-level defence. Despite their potency, FIAs typically require physical access and tailored fault models.

AI and ML-assisted attacks increasingly blur the line between classical and side-channel techniques. Attacks that combine logic block encryption analysis with neural networks, or that employ padding oracle exploitation enhanced by generative AI, illustrate the growing role of automation in key extraction. However, these approaches remain largely dependent on data labelling, training overhead, and generalization performance factors that currently limit their scalability across different hardware environments.

Quantum computing threats, while still largely theoretical, have begun to transition into concrete frameworks using simulation environments and hybrid quantum-classical models. Attacks based on Grover's algorithm, quantum amplitude amplification, and Simon's algorithm collectively demonstrate a consistent reduction in brute-force complexity and highlight specific vulnerabilities in mode-based AES instantiations such as AES-OTR. Nevertheless, these methods are contingent on the future availability of large-scale quantum hardware and remain infeasible in practice at present.

To systematically analyze recent advancements in AES attacks, we present two structured summaries. Table 1 categorizes key studies by attack type (side-channel, machine learning, or fault injection), their target (e.g., hardware/software implementation), and complexity—classified as low (O(n), O(n^2)), medium (O(2^(n/2))), or high (O(2^n))—along with their practical feasibility. This scaling highlights the computational tractability of attacks, aiding in risk assessment. Table 2 complements this by detailing data requirements, proposed countermeasures, key conclusions, and limitations of each study, emphasizing gaps in mitigation strategies. Together, these tables provide a consolidated framework for evaluating attack trends, computational trade-offs, and defensive priorities in AES security research.

Table 1.
Attacks Characteries

Paper (Year)	Attack Type	Target	Complexity	Practical or Theoretical?
Kuroda, et al. [9]	Side-channel (DL)	AES software with masking	Medium	Practical
Liu, et al. [15]	Side-channel (Frequency Throttling)	General AES implementations	High	Practical
Wang [14]	Side-channel (EM)	Masked AES	Medium	Practical
Fukuda, et al. [10]	Side-channel (DL, Multi-bit)	AES hardware	High	Practical
Zhong, et al. [21]	Machine learning aided key guessing	Logic block encryption	Medium	Practical
Pasquale Arpaia [22]	Side-channel (ML-based power)	AES hardware	Medium	Practical
Priya and Kapilamithran [23]	Padding Oracle / AI-based	AES CBC	Medium	Practical
Negabi, et al. [13]	Side-channel (Power Analysis + DL)	AES on ATmega328P	Medium	Practical
Daehyeon Bae, et al. [12]	Side-channel (DL)	Masked AES	Medium	Practical
Ghandali, et al. [11]	Side-channel (Power, DL)	AES-128	High	Practical
Swaminathan, et al. [25]	Side-channel (DL)	AES inner rounds	High	Practical
Kubota, et al. [26]	Side-channel (DL)	AES hardware	Medium	Practical
Tom´a ˇs Gerlich [27]	See-in-the-middle (DL)	AES	High	Theoretical
Dhanalakshmi [28]	DL Classification Attack	WSN Encryption	Medium	Practical
Xiaoyang Dong [29]	Classical + Quantum Rebound	AES-like Hashing	High	Theoretical
Chang, et al. [30]	Quantum Collision Forgery	AES-OTR	High	Theoretical
Jang, et al. [31]	Quantum Analysis	AES	Low	Theoretical
Cai, et al. [32]	Quantum Attack	1K-AES, PRINCE	Low	Theoretical
Mandal, et al. [33]	Quantum (Grover)	AES-based AEAD	Medium	Theoretical
Moiseevskiy [35]	Quantum-enhanced Classical	S-AES	Low	Theoretical
Sheikhpour, et al. [16]	Fault Injection	AES in IoT	Low	Practical
Pan, et al. [17]	DCA + Fault Injection	White-box AES	High	Practical
Ghosal, et al. [18]	Differential Fault Analysis	AES Hardware	Medium	Practical
Mestiri, et al. [20]	Fault Detection	AES Embedded Systems	Low	Practical

Table 2.
Attacks Implication

Paper (Year)	Data Requirements	Countermeasures	Key Conclusion	Limitation
Kuroda, et al. [9]	Side-channel traces	Masking (RSM)	DL attacks can break masked AES under certain settings	Depends on training data and leakage model
Liu, et al. [15]	Power profile timing	Constant frequency, detection systems	Frequency-based leakage is exploitable	Requires system-level access
Wang [14]	EM traces	EM shielding, masking	Even masked AES leaks through EM	Setup-dependent
Fukuda, et al. [10]	Power traces	Multi-bit masking	DL attacks can use multiple-bit labels effectively	Data-hungry and model sensitive
Zhong, et al. [21]	Plaintext-ciphertext pairs	Obfuscation, randomization	ML can speed up guessing attacks	Assumes labeled data
Pasquale Arpaia [22]	Power measurements	Noise injection	Uncertainty must be managed in ML attacks	Accuracy tradeoffs
Priya and Kapilamithran [23]	Access to padding oracle	Proper error handling	AI can exploit padding oracles	Requires oracle access
Negabi, et al. [13]	Power traces	Hiding, masking	DL aids in power analysis key extraction	Needs device access
Daehyeon Bae, et al. [12]	Power traces	Masking	DL models can recover keys from masked AES	Relies on training quality
Ghandali, et al. [11]	Power traces	Power equalization	K-TSVM improves side-channel key recovery	Computationally intensive
Swaminathan, et al. [25]	Power traces	Round randomization	DL can target AES rounds independently	Requires accurate labeling
Kubota, et al. [26]	Power traces	Hardware masking	DL models can exploit hardware leakages	Needs device-specific tuning
Tom´a ˇ s Gerlich [27]	Intermediate values	Layer separation	DL-SITM reveals internal AES state	Hard to generalize
Dhanalakshmi [28]	Encrypted traffic	Encryption variation	DL enhances traffic classification	Specific to WSN
Xiaoyang Dong [29]	Related-key differentials	Avoid related keys	Rebound attack still viable in hybrid settings	Not full AES
Chang, et al. [30]	Ciphertext pairs	Larger tag sizes	Quantum threats exist for authenticated encryption	Model-dependent
Jang, et al. [31]	Key space model	AES-256	AES security halves under Grover	Quantum model assumption
Cai, et al. [32]	Ciphertexts	High key length	Short key AES variants vulnerable	Quantum feasibility
Mandal, et al. [33]	Ciphertext queries	AES-256	Grover reduces AEAD scheme strength	Assumes large-scale quantum
Moiseevskiy [35]	Known plaintext	Quantum-resistant ciphers	Quantum models accelerate classical attacks	Toy cipher (S-AES)
Sheikhpour, et al. [16]	Faulty ciphertexts	Redundancy	Low-cost hardware can resist fault attacks	Limited scalability
Pan, et al. [17]	Traces + faults	Obfuscation	Fault + DCA effective against white-box AES	Resource-intensive
Ghosal, et al. [18]	Faulty ciphertext	Fault-tolerant design	Secure hardware can tolerate faults	Extra resource cost

Mestiri, et al. [20]	System logs	Real-time error monitoring	Detects anomalies in AES execution	May not block attacks

The analysis reveals that most of recent AES attacks are practical in nature, with 17 categorized as practical and only 7 as theoretical as shown in figure 1. This indicates a strong trend in the cryptanalysis community toward implementing and validating attack methods in real-world or simulated environments. The prevalence of practical attacks highlights the importance of considering applied vulnerabilities in AES, especially in constrained or specific use-case scenarios like IoT and embedded systems.

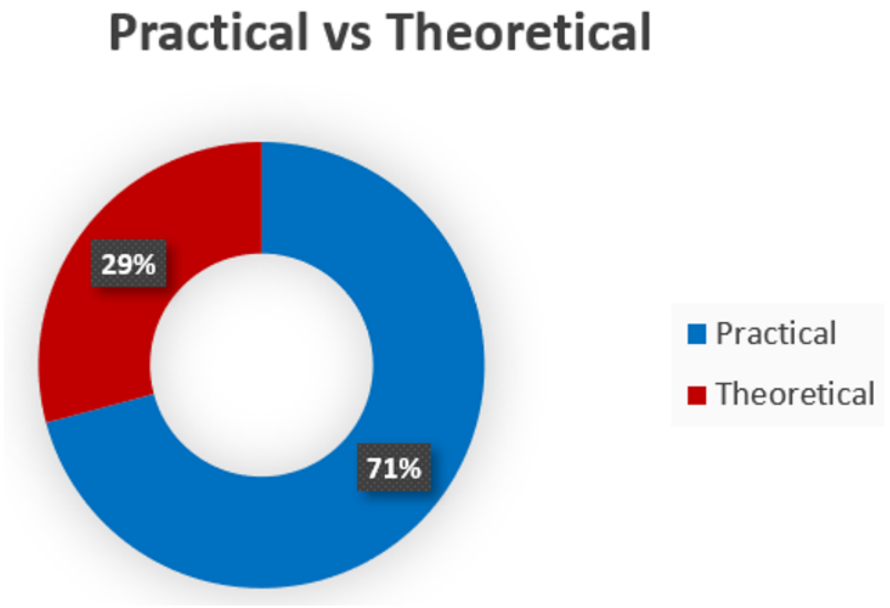

Figure 1.
Practical Attacks vs Theoretical Attacks.

Figure 2 analyzed attacks, 5 were classified as low complexity, 11 as medium complexity, and 8 as high complexity. The predominance of medium complexity attacks suggests that most breakthroughs strike a balance between feasibility and sophistication, making them significant enough to warrant attention yet accessible enough to be reproduced or extended. The relatively high number of high-complexity attacks also reflects ongoing efforts to push the boundaries of AES security through advanced techniques such as quantum cryptanalysis and fault injections.

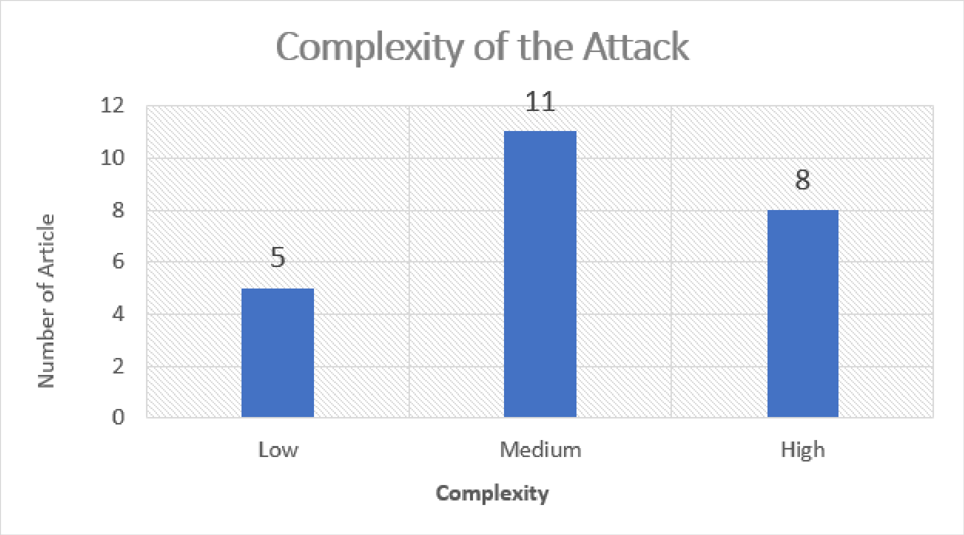

Figure 2.
Number of articles for each complexity class

When examining attack complexity over the years in figure 3, a diverse distribution is observed across low, medium, and high complexity categories each year. This suggests that no single trend dominates the field annually; instead, researchers explore vulnerabilities across all complexity levels. However, a gradual increase in medium and high complexity attacks in recent years may reflect the maturing of cryptanalytic methods and the growing use of advanced computing resources, including AI and quantum models, to challenge AES resilience.

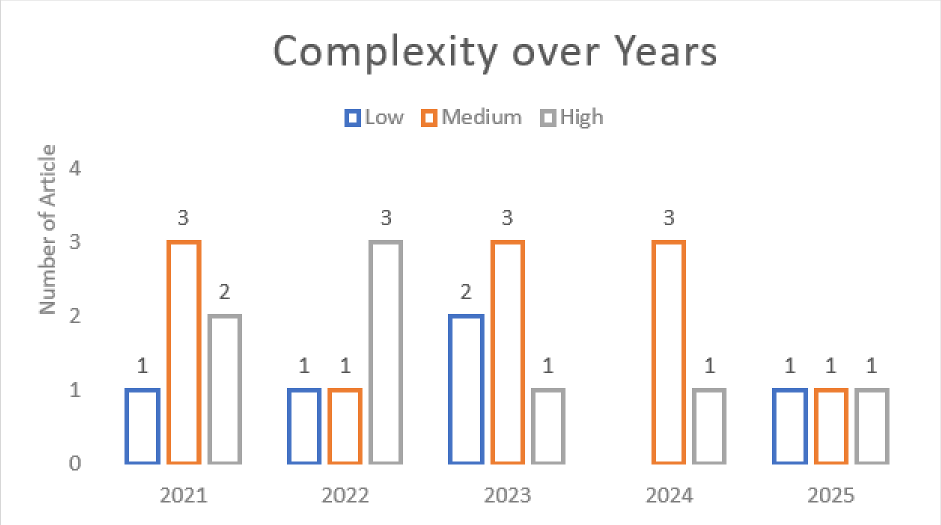

Figure 3.
Number of articles for each year classify by their complexity

Based on the Table 2, the following limitations are observed across attack categories:

- Many practical attacks require extensive side-channel traces (≥50K) or faults (≥100), which are not trivial to obtain in field conditions.

- Attack complexity, especially for deep learning and quantum variants, is non-trivial—training models or simulating quantum algorithms incurs substantial computational cost.
- Countermeasures like masking, redundancy, and noise injection continue to challenge the reproducibility and robustness of many attacks.
- Theoretical attacks often make idealized assumptions about attacker capabilities or fault models that do not hold in constrained, real-world deployments.
- Few attacks are demonstrated across diverse platforms (e.g., FPGA, ASIC, microcontrollers), limiting generalization across architectures.
- While attacks continue to evolve, they simultaneously illuminate vulnerabilities and catalyse the development of more robust AES implementations. This dynamic arms race is critical to sustaining the long-term viability of AES in both legacy and forward-looking systems.

5. Conclusion and Future Work

AES has remained a cornerstone of modern cryptography for over two decades, trusted for its simplicity, speed, and resistance to classical cryptanalysis. However, the continued advancement of attack methodologies especially those employing side-channel leakage, machine learning, and quantum computing demands ongoing scrutiny and innovation in its implementation.

This review synthesized recent advances in AES cryptanalysis from 2021 to 2025, categorizing them across four major domains: side-channel attacks, fault injection technique, machine learning-assisted techniques and quantum computing threats. Each attack type was evaluated based on its complexity, data requirements, countermeasures, and practical feasibility. The comparative analysis reveals that many attacks are now capable of targeting full-round AES-128 implementations under realistic settings, particularly in embedded and IoT contexts.

Looking forward, several avenues for future research are evident:

- Scalable and transferable machine learning models that generalize across devices and implementations.
- Lightweight but effective countermeasures for AES in resource-constrained environments, especially for wearable and mobile technologies.
- Robust benchmarking frameworks and datasets to enable reproducible evaluation of side-channel and fault attacks.
- Post-quantum symmetric design adaptations, such as increased key lengths or hybrid modes, to resist quadratic quantum speedups.
- Integrated co-design approaches, where hardware, firmware, and cryptographic logic are developed jointly with fault and leakage resilience in mind.

Ultimately, while AES remains secure under its original design assumptions, its real-world security is increasingly dependent on implementation details. This review underscores the necessity of continuously evolving both offensive and defensive strategies to preserve the integrity of AES-based systems in a rapidly changing technological landscape.

Transparency:

The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.

Copyright:

References

[1] A. Altigani, S. Hasan, B. Barry, S. Naserelden, M. A. Elsadig, and H. T. Elshoush, "A polymorphic advanced encryption standard–a novel approach," *IEEE Access*, vol. 9, pp. 20191-20207, 2021. https://doi.org/10.1109/ACCESS.2021.3055036

[2] S. Koley and A. Kumar, *Enhancing video steganography security for cross-platform applications: A focus on high-definition formats and streaming environments. In Next-Generation Systems and Secure Computing*. United Kingdom: Wiley, 2025.

[3] A. Altigani, S. Hasan, B. Barry, and S. M. Shamsuddin, "Key-dependent advanced encryption standard," in *2018 International Conference on Computer, Control, Electrical, and Electronics Engineering (ICCCEEE), IEEE*, 2018, pp. 1-5.

[4] A. Flórez-Gutiérrez, L. Grassi, G. Leander, F. Sibleyras, and Y. Todo, "General practical cryptanalysis of the sum of round-reduced block ciphers and ZIP-AES," in *International Conference on the Theory and Application of Cryptology and Information Security*, 2025: Springer, pp. 280-311.

[5] A. Altigani, M. R. Ghalib, H. T. Elshoush, M. A. Elsadig, N. Alrefai, and S. Naserelden, "The advanced encryption standard with a dynamic S-box using the RC4 key schedule," in *2024 International Conference on Electrical, Computer and Energy Technologies, IEEE*, 2024, pp. 1-7.

[6] P. R. Sekhar Reddy *et al.*, "Adaptive and scalable cloud data sharing framework with quantum-resistant security, decentralized auditing, and machine learning-based threat detection," *International Journal of Advanced Computer Science & Applications*, vol. 16, no. 2, p. 1042, 2025. https://doi.org/10.14569/IJACSA.2025.01602103

[7] N. U. Ain, S.-S. Ahmadpour, N. J. Navimipour, E. Diakina, and S. R. Kassa, "Secure quantum-based adder design for protecting machine learning systems against side-channel attacks," *Applied Soft Computing*, vol. 169, p. 112554, 2025. https://doi.org/10.1016/j.asoc.2024.112554

[8] A. Altigani, S. Hasan, and B. Barry, "The need for polymorphic encryption algorithms: A review paper," *Journal of Theoretical and Applied Information Technology*, vol. 98, no. 3, pp. 360-377, 2020.

[9] K. Kuroda, Y. Fukuda, K. Yoshida, and T. Fujino, "Practical aspects on non-profiled deep-learning side-channel attacks against AES software implementation with two types of masking countermeasures including RSM," in *Proceedings of the 5th Workshop on Attacks and Solutions in Hardware Security*, 2021, pp. 29-40, doi: https://doi.org/10.1145/3474376.3487285.

[10] Y. Fukuda, K. Yoshida, and T. Fujino, "Multi-bit DDLA: Non-profiled deep learning side-channel attacks using multi-bit label against hardware-implemented AES," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E108.A, no. 3, pp. 215-226, 2025. https://doi.org/10.1587/transfun.2024CIP0009

[11] S. Ghandali, S. Ghandali, and S. Tehranipoor, "Deep K-TSVM: A novel profiled power side-channel attack on AES-128," *IEEE Access*, vol. 9, pp. 136448-136458, 2021. https://doi.org/10.1109/ACCESS.2021.3111234

[12] D. B. Daehyeon Bae, J. H. Daehyeon Bae, and J. H. Jongbae Hwang, "Deep learning-based attacks on masked AES implementation," *Internet Technology Journal*, vol. 23, no. 4, pp. 897-902, 2022. https://doi.org/10.53106/160792642022072304024

[13] I. Negabi, S. Ait El Asri, S. El Adib, and N. Raissouni, "Deep learning-based power analysis attack for extracting AES keys on ATmega328P microcontroller," *Arabian Journal for Science and Engineering*, vol. 49, no. 3, pp. 4197-4208, 2023. https://doi.org/10.1007/s13369-023-08341-3

[14] H. Wang, "Amplitude-modulated EM side-channel attack on provably secure masked AES," *Journal of Cryptographic Engineering*, vol. 14, no. 3, pp. 537-549, 2024. https://doi.org/10.1007/s13389-024-00347-3

[15] C. Liu, A. Chakraborty, N. Chawla, and N. Roggel, "Frequency throttling side-channel attack," in *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security (CCS '22), ACM*, 2022, pp. 1977-1991, doi: 10.1145/3548606.3560682.

[16] S. Sheikhpour, S.-B. Ko, and A. Mahani, "A low cost fault-attack resilient AES for IoT applications," *Microelectronics Reliability*, vol. 123, p. 114202, 2021. https://doi.org/10.1016/j.microrel.2021.114202

[17] Y. Pan *et al.*, "A DCA attack scheme and a genetic algorithm-based intelligent fault injection algorithm for white-box AES implementation," in *Proceedings of the 2023 IEEE Asian Solid-State Circuits Conference (ASSCC), IEEE*, 2023, pp. 72-76, doi: https://doi.org/10.1109/ASID60355.2023.10426571.

[18] A. K. Ghosal, A. Sardar, and D. R. Chowdhury, "Differential fault analysis attack-tolerant hardware implementation of AES," *The Journal of Supercomputing*, vol. 80, no. 4, pp. 4648-4681, 2023. https://doi.org/10.1007/s11227-023-05632-2

[19] R. Anand, T. Isobe, A. K. Kundu, M. Rahman, and S. Suryawanshi, "Differential fault attack on AES-based encryption schemes: Application to B5G/6G ciphers—Rocca, Rocca-S and AEGIS," *Journal of Cryptographic Engineering*, vol. 14, no. 4, pp. 595-607, 2024. https://doi.org/10.1007/s13389-024-00360-6

[20] H. Mestiri, I. Barraj, and M. Machhout, "Innovative fault detection for AES in embedded systems: Advancing resilient and sustainable digital security," *Engineering, Technology & Applied Science Research*, vol. 15, no. 2, pp. 20660-20667, 2025. https://doi.org/10.48084/etasr.9852

[21] Y. Zhong, J.-H. Feng, X.-X. Cui, and X.-L. Cui, "Machine Learning Aided Key-Guessing Attack Paradigm Against Logic Block Encryption," *Journal of Computer Science and Technology*, vol. 36, no. 5, pp. 1102-1117, 2021. https://doi.org/10.1007/s11390-021-0846-6

[22] F. C. Pasquale Arpaia, Antonella Cioffi , Antonio Esposito and Francesco Isgrò, "Uncertainty analysis in cryptographic key recovery for machine learning-based power measurements attacks," *IEEE Communications Surveys & Tutorials*, vol. 72, pp. 1-8, 2023. https://doi.org/10.1109/COMST.2023.3245678

[23] L. Priya and S. Kapilamithran, "Exploiting AES encryption vulnerabilities through padding oracle attacks and generative AI techniques," in *Proceedings of the 2024 IEEE International Conference on Smart Technologies and Management (i-SMAC), IEEE*, 2024, pp. 682-689, doi: https://doi.org/10.1109/i-smac61858.2024.10714697.

[24] J. Ahn, R. Hussain, K. Kang, and J. Son, "Exploring encryption algorithms and network protocols: A comprehensive survey of threats and vulnerabilities," *IEEE Communications Surveys & Tutorials*, vol. 27, no. 2, pp. 1-10, 2025. https://doi.org/10.1109/comst.2025.3526605

[25] S. Swaminathan, Ł. Chmielewski, G. Perin, and S. Picek, "Deep learning-based side-channel analysis against aes inner rounds," in *International Conference on Applied Cryptography and Network Security, Springer*, 2022, pp. 165-182.

[26] 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. https://doi.org/10.1016/j.micpro.2021.103383

[27] J. B. Tom´aˇs Gerlich, Pavel Sikora, Zdenˇek Martin´asek, Aron Gohr, Anubhab Baksi and Xiaolu Hou, "DL-SITM: Deep learning-based see-in-the-middle attack on AES," *IACR Cryptology ePrint Archive*, p. 1389, 2024. https://eprint.iacr.org/2024/1389

[28] N. Dhanalakshmi, "Unmasking encryption effects and modified Deep learning approaches for attack classification in WSN," *Expert Systems with Applications*, vol. 266, p. 126163, 2025. https://doi.org/10.1016/j.eswa.2024.126163

[29] Z. Z. Xiaoyang Dong, Siwei Sun, Congming Wei Xiaoyun Wang and Lei Hu, "Automatic classical and quantum rebound attacks on AES-like hashing by exploiting related-key differentials," in *Advances in Cryptology–ASIACRYPT 2021: 27th International Conference on the Theory and Application of Cryptology and Information Security, Singapore, December 6–10, 2021, Proceedings, Part I 27, Springer International Publishing*, 2021, pp. 241-271.

[30] L. Chang, Y. Wei, X. Wang, and X. Pan, "Collision forgery attack on the AES-OTR algorithm under quantum computing," *Symmetry*, vol. 14, no. 7, p. 1434, 2022. https://doi.org/10.3390/sym14071434

[31] K. Jang, A. Baksi, H. Kim, G. Song, H. Seo, and A. Chattopadhyay, "Quantum analysis of AES," Cryptology ePrint Archive, Paper No. 683, 2022.

[32] B.-B. Cai, Y. Wu, J. Dong, S.-J. Qin, F. Gao, and Q.-Y. Wen, "Quantum attacks on 1K-AES and PRINCE," *The Computer Journal*, vol. 66, no. 5, pp. 1102-1110, 2023. https://doi.org/10.1093/comjnl/bxab216

[33] S. Mandal, R. Anand, M. Rahman, S. Sarkar, and T. Isobe, "Implementing Grover's on AES-based AEAD schemes," *Scientific Reports*, vol. 14, no. 1, p. 21105, Sep 10 2024. https://doi.org/10.1038/s41598-024-69188-8

[34] A. Baksi and K. Jang, "Quantum analysis of AES, implementation and analysis of ciphers in quantum computing," in *Applications and Techniques in Information Security: 14th International Conference, ATIS 2024, Tamil Nadu, India, November 22–24, 2024, Proceedings,Springer Singapore*, 2024, pp. 1–13.

[35] A. Moiseevskiy, "Quantum-enhanced symmetric cryptanalysis for S-AES," *arXiv preprint arXiv:2304.05380*, 2023. https://arxiv.org/abs/2304.05380