Advanced Circuit Techniques for Secure Analog Neural Implementations

Atika Nishat, Areej Mustafa

Department of Information Technology, University of Gujrat, Pakistan

Department of Information Technology, University of Gujrat, Pakistan

Abstract:

With the rise of artificial intelligence (AI) and neural network models in various sectors, securing the hardware implementations of neural networks has become increasingly critical. Analog neural networks, known for their power efficiency and high-speed computations, offer a promising solution for AI-based applications. However, the integration of neural processing units (NPUs) in analog circuits raises significant concerns about security, privacy, and integrity. This paper explores advanced circuit techniques designed to enhance the security of analog neural network implementations. These methods address vulnerabilities such as signal leakage, circuit manipulation, and adversarial attacks, which can compromise the performance and integrity of neural networks. The focus is on circuit-level innovations, including secure training methods, hardware obfuscation, side-channel attack mitigation, and tamper detection. By leveraging advanced design techniques, this work seeks to pave the way for more secure and reliable analog hardware for neural network applications, particularly in critical areas such as finance, healthcare, and defense.

Keywords: Analog neural networks, hardware security, signal leakage, adversarial attacks, tamper detection, circuit obfuscation, neural network hardware, analog security techniques

I. Introduction

Analog neural networks (ANNs) are gaining attention due to their potential advantages in processing efficiency and speed, particularly when compared to digital alternatives. These networks rely on continuous signals rather than discrete ones, enabling faster calculations with reduced energy consumption [1]. Despite these advantages, there are significant concerns

regarding the security of analog circuits that implement neural networks. Traditional digital neural networks have already been the subject of extensive research in secure hardware implementations, but the challenges presented by analog systems are less well understood. Security in analog neural implementations is critical as AI continues to evolve and permeate sensitive areas such as autonomous vehicles, military systems, and personal healthcare devices. Analog circuits, while offering performance benefits, are inherently vulnerable to various types of attacks, such as side-channel attacks and tampering [2]. Analog systems are also susceptible to issues such as signal leakage and unauthorized access, which can undermine the confidentiality and integrity of the data being processed [3]. In this paper, we explore advanced circuit techniques designed to mitigate these security risks in analog neural network implementations. These include methods for enhancing resistance to side-channel attacks, implementing circuit obfuscation techniques, and employing tamper detection mechanisms. These innovations aim to secure the data flow within neural networks and protect them from malicious interventions [4].

II. Circuit Design Challenges in Analog Neural Networks

Analog circuits present unique challenges when implementing secure neural networks. One of the key difficulties lies in the inherent complexity of analog signal processing, which makes it more susceptible to external interference [5]. Unlike digital systems, which can be isolated through encryption and other digital security methods, analog systems are often directly influenced by environmental factors such as electromagnetic radiation, power consumption fluctuations, and temperature variations [6]. The continuous nature of analog signals makes it difficult to apply traditional cryptographic techniques used in digital systems. For instance, analog circuits cannot easily mask or encrypt signals during processing, making them vulnerable to eavesdropping. This exposure allows attackers to gain insights into the data being processed and exploit these weaknesses for malicious purposes [7].

Moreover, the physical nature of analog circuits allows for the possibility of hardware manipulation. Attackers can alter the circuit's components or introduce faults to manipulate the output of the neural network, potentially causing incorrect predictions or classifications. The challenge here lies in designing circuits that are both resilient to these attacks and capable of detecting tampering in real-time [8]. Another challenge is the limited ability to implement

conventional error detection and correction mechanisms in analog hardware. In digital systems, error correction codes and redundant circuits can be used to ensure the integrity of data. However, in analog systems, these methods are less effective due to the continuous nature of the signals [9]. Designing robust analog circuits that can detect and respond to errors without compromising performance is a critical task for secure neural network implementations [10].

III. Secure Training Methods in Analog Neural Networks

Training neural networks is a process that requires exposing the system to large amounts of data, which can be susceptible to interception or manipulation. Secure training methods are therefore essential to ensure that the learned models remain confidential and resistant to adversarial interventions. In analog systems, where signals are continuous and often lack the encryption features of digital systems, there is an increased risk that an attacker could intercept or manipulate training data [11]. Techniques like data masking and secure multi-party computation (SMPC) are gaining traction in digital systems, but adapting these techniques to analog systems presents a significant challenge. Advanced analog circuits can be designed to perform encrypted computations during the training phase.

For example, homomorphic encryption, which allows computations on encrypted data, can be implemented in analog hardware to ensure that sensitive information remains protected during processing. However, this requires the development of specialized circuits that can efficiently handle encrypted analog signals without introducing significant overhead. Another approach to secure training is the use of adversarial training. This involves deliberately introducing perturbations into the training process to make the neural network more robust against attacks. In the context of analog circuits, this could involve designing circuits that are inherently resistant to small variations in input signals, ensuring that adversarial attacks do not have a significant impact on the network's learning process.

Additionally, implementing techniques such as differential privacy during training can help ensure that the network does not leak sensitive information about the individual data points in the training set. Differential privacy ensures that the output of the neural network is statistically

indistinguishable, regardless of whether a particular individual's data is included in the training set, thus protecting privacy during the training phase.

IV. Side-Channel Attack Mitigation in Analog Neural Networks

Side-channel attacks exploit the physical characteristics of a circuit to gain insights into its internal workings. In the case of analog circuits, side-channel attacks often focus on analyzing variations in power consumption, electromagnetic emissions, or even acoustic signals to extract confidential information. Analog neural networks are particularly vulnerable to such attacks due to the continuous nature of the signals involved. To mitigate the risk of side-channel attacks, one approach is to introduce noise or jitter into the analog circuits. By intentionally adding small, random fluctuations to power consumption or signal processing, attackers are prevented from accurately measuring these characteristics and extracting useful information. This approach, while effective in obscuring the internal workings of the circuit, must be carefully calibrated to avoid negatively impacting the performance of the neural network.

Another strategy involves designing the circuit to be resistant to electromagnetic interference. Shielding techniques, such as the use of Faraday cages or other conductive materials, can be employed to block unwanted emissions from the circuit. This can prevent attackers from using external equipment to monitor the circuit's electromagnetic output and gain access to sensitive information. Power analysis attacks, in which attackers measure the power consumption of a circuit to infer information about its internal state, can be mitigated through techniques like dynamic voltage and frequency scaling (DVFS). By varying the voltage and frequency of the circuit's operations in a non-deterministic manner, attackers are unable to correlate power consumption patterns with specific operations or data being processed.

Another key area for side-channel attack mitigation is the use of secure analog-to-digital conversion (ADC) techniques. Since analog signals must eventually be digitized for processing, ensuring that the ADC does not introduce vulnerabilities is critical. By employing secure ADC designs that minimize leakage and maintain the integrity of the data being converted, the overall security of the analog neural network can be significantly improved.

V. Circuit Obfuscation for Security in Analog Neural Networks

Circuit obfuscation is a technique used to make it more difficult for attackers to reverse-engineer or manipulate hardware designs. In analog neural networks, circuit obfuscation can be applied to prevent the unauthorized understanding or modification of the neural network's structure and functionality. Obfuscation methods can involve altering the design of the analog circuit in ways that make it challenging to predict the network's behavior. For example, circuit-level transformations, such as the addition of redundant components or the introduction of complex feedback loops, can make it harder for an attacker to understand the logic behind the network's operations [12].

Another approach to obfuscation is to use polymorphic circuits, which change their behavior over time or in response to external stimuli. By altering the circuit's functionality dynamically, attackers are unable to determine the precise configuration of the network at any given moment, thus preventing them from exploiting vulnerabilities. Additionally, hardware encryption can be integrated into the circuit's design to protect the intellectual property of the neural network. By encrypting key components of the design, such as the weights of the neural network or the configuration of individual processing units, the network can be protected from reverse-engineering attempts.

Obfuscation techniques must be carefully balanced to avoid significant overhead or performance degradation. Excessive obfuscation can increase the complexity of the circuit, which may lead to higher power consumption or slower processing speeds. Therefore, an optimal approach must be found that ensures both security and performance are maintained.

VI. Tamper Detection and Response Mechanisms

Tamper detection is crucial in ensuring the integrity of analog neural networks. Since analog circuits are prone to physical attacks, such as hardware modifications or signal manipulation, detecting these tampering attempts in real time is essential for preventing damage or data breaches. One effective approach is the use of built-in sensors that can monitor the physical conditions of the circuit, such as temperature, voltage, and current. Deviations from normal

operating conditions can indicate potential tampering, prompting an immediate response, such as shutting down the circuit or alerting the system to the potential threat.

Tamper detection circuits can also incorporate self-test mechanisms that check the integrity of key components. For example, the weights of the neural network could be periodically verified against expected values to ensure that they have not been altered. If discrepancies are detected, the system can trigger a re-training process or alert the operator to investigate further.

In addition to passive detection, active countermeasures can be employed to respond to tampering attempts. For example, if tampering is detected, the system could initiate a process to erase sensitive data or disrupt the functioning of the neural network. This makes it more difficult for attackers to extract useful information even if they successfully alter the hardware. These tamper detection and response mechanisms are essential for ensuring the continued security of analog neural networks, particularly in high-stakes environments where the consequences of a breach could be severe.

Conclusion

As the adoption of analog neural networks continues to grow, ensuring their security becomes increasingly important. The unique characteristics of analog circuits such as their continuous nature and susceptibility to physical attacks present significant challenges for securing these systems. However, by leveraging advanced circuit techniques, it is possible to mitigate many of these risks and enhance the integrity and privacy of analog neural networks. The application of secure training methods, side-channel attack mitigation, circuit obfuscation, and tamper detection offers promising solutions to these challenges. By integrating these techniques into analog neural network designs, it is possible to create hardware systems that are not only efficient and fast but also resilient to malicious attacks and unauthorized manipulation. As the field of secure analog neural networks evolves, further research is needed to refine these techniques and develop new approaches to address emerging threats. However, the advancements outlined in this paper provide a solid foundation for the continued development of secure analog neural network hardware, with the potential to revolutionize AI applications across a range of industries.

REFERENCES:

[1] M. R. Abdelhamid, R. Chen, J. Cho, A. P. Chandrakasan, and F. Adib, "Self-reconfigurable micro-implants for cross-tissue wireless and batteryless connectivity," in *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking*, 2020, pp. 1-14.

[2] L. De Marinis *et al.*, "A codesigned integrated photonic electronic neuron," *IEEE Journal of Quantum Electronics,* vol. 58, no. 5, pp. 1-10, 2022.

[3] R. Chen, A. Chandrakasan, and H. Lee, "Direct Hybrid Encoding for Signed Expressions SAR ADC for Analog Neural Networks," *Circuits & Systems for Communications, IoT, and Machine Learning,* p. 23, 2021.

[4] R. Chen, H. Kung, A. Chandrakasan, and H. Lee, "A Bit-level Sparsity-aware SAR ADC with Direct Hybrid Encoding for Signed Expressions Leveraging Algorithm-circuit Co-design," *Circuits, Systems, and Power Electronics,* p. 23, 2022.

[5] W. Haensch, T. Gokmen, and R. Puri, "The next generation of deep learning hardware: Analog computing," *Proceedings of the IEEE,* vol. 107, no. 1, pp. 108-122, 2018.

[6] R. Chen, "Activity-Scaling SAR with Direct Hybrid Encoding for Signed Expressions for AIoT Applications," Massachusetts Institute of Technology, 2021.

[7] R. Chen, H. Kung, A. Chandrakasan, and H.-S. Lee, "A bit-level sparsity-aware SAR ADC with direct hybrid encoding for signed expressions for AIoT applications," in *Proceedings of the ACM/IEEE International Symposium on Low Power Electronics and Design*, 2022, pp. 1-6.

[8] R. Chen, "Analog-to-Digital Converters for Secure and Emerging AIoT Applications," Massachusetts Institute of Technology, 2023.

[9] K. J. Lee, "Architecture of neural processing unit for deep neural networks," in *Advances in Computers*, vol. 122: Elsevier, 2021, pp. 217-245.

[10] R. Chen, A. Chandrakasan, and H.-S. Lee, "Sniff-sar: A 9.8 fj/c.-s 12b secure adc with detectiondriven protection against power and em side-channel attack," in *2023 IEEE Custom Integrated Circuits Conference (CICC)*, 2023: IEEE, pp. 1-2.

[11] R. Chen, H. Wang, A. Chandrakasan, and H.-S. Lee, "RaM-SAR: a low energy and area overhead, 11.3 fj/conv.-step 12b 25ms/s secure random-mapping SAR ADC with power and EM side-channel attack resilience," in *2022 IEEE Symposium on VLSI Technology and Circuits (VLSI Technology and Circuits)*, 2022: IEEE, pp. 94-95.

[12] M. Musisi-Nkambwe, S. Afshari, H. Barnaby, M. Kozicki, and I. S. Esqueda, "The viability of analog-based accelerators for neuromorphic computing: a survey," *Neuromorphic Computing and Engineering,* vol. 1, no. 1, p. 012001, 2021.