Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

EM Side Channels for Obfuscation-Aware Malware Detection on Intel: Packers and Virtualization

Sergio López-Flores
University of Malaga
Antonio Muñoz
anto@uma.es
University of Malaga

Research Article

Keywords: Malware detection, Obfuscation, Electromagnetic side-channel analysis, Intel processors, Code virtualization

Posted Date: March 5th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8925591/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

EM Side Channels for Obfuscation-Aware Malware Detection on Intel: Packers and Virtualization

Sergio Lopez-Flores[a], Antonio Muñoz[a,*]

[a]*Network, Information and Computer Security (NICS) Lab, University of Malaga, Spain*

Abstract

Executable packing and code virtualization are widely used to impede malware analysis, weakening both static inspection and in-host dynamic monitoring. This paper evaluates whether near-field electromagnetic (EM) leakage can expose these obfuscation layers on commodity Intel desktop hardware without any software instrumentation.

We study two binary tasks, packed vs. unpacked and virtualized vs. non-virtualized, using paired variants derived from three Linux malware families, with executions interleaved with realistic benign background activity. Two acquisition and representation strategies are compared: a high-fidelity PicoScope-based chain processed through STFT spectrograms with NICV-guided frequency selection, and *EM-SENSE*, a low-cost Arduino-based prototype that operates at low sampling rates and uses direct time-domain encoding.

Across consistent experimental splits and a shared model suite, EM emissions retain discriminative structure for both packing and virtualization. High-fidelity measurements reach up to 99% accuracy for packing and 94% for virtualization, while the low-cost prototype remains above 90% on both tasks. Classical LDA-based pipelines consistently outperform deep architectures, indicating that carefully engineered representations are more effective than end-to-end learning under the studied acquisition conditions.

These results support EM side-channel sensing, including low-cost hardware, as a practical, complementary signal for detecting modern malware obfuscation.

Keywords: Malware detection, Obfuscation, Electromagnetic side-channel analysis, Intel processors, Code virtualization.

1. Introduction

The evolution of malware has transformed it from experimental software created for intellectual curiosity into a major cybersecurity threat capable of disrupting economies, governments, and individuals around the world. Modern malware encompasses a wide spectrum of threats, ranging from ransomware and spyware to kernel-level rootkits and advanced persistent threats (APTs). Unlike early generations, which were relatively simple and often easy to detect, contemporary malware increasingly relies on sophisticated evasion mechanisms designed to bypass traditional defensive solutions Sutherland (2013).

A defining characteristic of modern malware is the widespread adoption of code obfuscation techniques, particularly packing and code virtualization. These techniques conceal malicious payloads, alter control flow, and obscure runtime behavior, severely limiting the effectiveness of both static and dynamic analysis. Packers compress or encrypt executables to hide malicious logic, while virtualization-based protectors translate original code into custom bytecode executed by an embedded virtual machine, effectively breaking conventional analysis workflows.

Beyond these technical advancements, the malware ecosystem itself has undergone a profound structural transformation. The emergence of Malware-as-a-Service (MaaS) has industrialized malware development, enabling the large-scale production, customization, and deployment of malicious software with built-in obfuscation, polymorphism, and anti-analysis features by default Patsakis et al. (2024). This service-oriented model dramatically accelerates the proliferation of evasive malware variants and significantly reduces the effectiveness of conventional software-based defenses, even against relatively unsophisticated threat actors.

Traditional malware detection techniques are increasingly limited in this context. Static analysis struggles when executables are packed or virtualized, while dynamic analysis can be actively evaded through environment fingerprinting, sandbox-awareness mechanisms, or delayed execution strategies. Although antivirus systems, intrusion detection solutions, and heuristic-based approaches have evolved considerably, the rapid pace of malware adaptation continues to allow many variants to evade detection. These limitations underscore the urgent need for alternative detection approaches that operate outside the software execution stack and cannot be influenced by malware behavior.

One promising direction is the exploitation of physical

*Corresponding author
Email address: anto@uma.es (Antonio Muñoz)

side channels of computing systems. In particular, electromagnetic (EM) emanations generated during program execution provide a non-intrusive and software-independent view of system activity. Because these emissions originate from the physical behavior of electronic components, they cannot be suppressed, hidden, or tampered with by software-level evasion techniques. EM-based monitoring therefore offers inherent robustness against packing, virtualization, and anti-instrumentation defenses.

While prior research has demonstrated the feasibility of EM side-channel analysis for malware detection Pham et al. (2021), important gaps remain. Most existing studies rely on expensive laboratory-grade equipment, focus primarily on ARM-based or embedded platforms, and often emphasize anomaly detection rather than explicit identification of obfuscation techniques. In contrast, Intel desktop processors, widely deployed in real-world environments, remain comparatively underexplored.

This work addresses these gaps by evaluating EM-based detection of malware obfuscation techniques on Intel CPUs, with a specific focus on binary classification of packed and virtualized malware. We conduct a systematic comparison between a professional PicoScope 5000 series acquisition setup and *EM-SENSE*, a low-cost, Arduino-based EM acquisition prototype designed for accessibility and reproducibility. By analyzing how acquisition fidelity impacts detection performance, this study provides practical insights into the feasibility of low-cost EM-based malware detection.

The main contributions of this work are summarized as follows:

1. *Obfuscation-Aware Malware Detection on Intel Architectures*: We demonstrate that EM side-channel analysis can reliably detect packing and virtualization techniques on Intel desktop processors.
2. *Comparison of Low-Cost and Professional EM Acquisition Hardware*: We quantify the trade-offs between signal fidelity, cost, and classification performance by comparing *EM-SENSE* with a laboratory-grade PicoScope setup.
3. *Robust Analysis Against Evasion Techniques*: We evaluate how EM-based methods perform against common malware evasion techniques, such as code packing and virtualization, demonstrating their resilience compared to software-based monitoring.
4. *Reproducible EM-Based Detection Pipeline*: We present a complete acquisition, preprocessing, and classification framework based on STFT and NICV feature selection, and release all datasets and scripts for reproducibility.

The rest of the paper is organized as follows: Section 2 reviews related work on EM-based malware detection. Section 3 describes the EM acquisition setup using both *EM-Sense* and the Picoscope 5000. Section 4 outlines the processing and classification framework. Section 5 presents the experimental scenarios. Section 6 compares the outcomes of both acquisition approaches. Section 7 analyzes the implications of the findings, and Section 8 concludes with final remarks and future work directions.

2. State-of-the-Art

Early research on hardware-assisted malware detection explored power consumption as a side-channel indicator. Clark et al. (2013) demonstrated that monitoring power traces could reveal malicious activity in embedded medical devices, although the scope of their study was limited to specific embedded platforms.

Building on this foundation, Khan et al. Khan et al. (2019b) showed that electromagnetic side-channel signals analyzed with Multilayer Perceptrons (MLPs) could effectively differentiate between benign and malicious behaviors in embedded systems. Subsequent work emphasized anomaly-based detection, where deviations from baseline EM patterns are interpreted as indicators of compromise. Representative examples include IDEA Khan et al. (2019a) and REMOTE Sehatbakhsh et al. (2019), both of which leveraged EM emissions to detect abnormal execution behavior.

Nazari et al. Nazari et al. (2017) combined Short-Time Fourier Transform (STFT) analysis with statistical tests such as the Kolmogorov–Smirnov test Berger and Zhou (2014) to detect execution anomalies in embedded devices. Riley et al. Riley et al. (2019) demonstrated that EM probes combined with Support Vector Machines (SVMs) could infer register-level activity, with deviations interpreted as attack signals. Wang et al. Wang et al. (2018) further expanded the application of machine learning to hardware side channels by employing autoencoders, Long Short-Term Memory (LSTM) networks, and MLPs for anomaly detection in side-channel traces.

More recent studies have explored the use of deep learning and time–frequency representations for malware detection using side-channel information. Ding et al. Ding et al. (2020) employed Convolutional Neural Networks (CNNs) to analyze power traces for IoT malware detection, while Chawla et al. Chawla et al. (2021) investigated wavelet-based feature extraction combined with machine learning for EM-based malware classification.

A significant contribution was presented by Pham et al. Pham et al. (2021), who demonstrated that EM side-channel emissions preserve discriminative information even in the presence of advanced obfuscation techniques such as packing and virtualization. However, their experiments focused on ARM-based IoT platforms and relied on high-end laboratory-grade oscilloscopes and probes, limiting reproducibility and accessibility.

Beyond EM-based approaches, other hardware-assisted detection mechanisms have been explored on Intel architectures, such as Performance Monitoring Units (PMUs) Demme et al. (2013). While promising, PMU-based meth-

Table 1: Definitions

Acronym	Definition	Acronym	Definition
EM	Electromagnetic	LDA	Linear Discriminant Analysis
IoT	Internet of Things	NB	Naive Bayes
ARM	Advanced RISC Machine	SVM	Support Vector Machines
DDoS	Distributed Denial-of-Service	RF	Random Forest
NICV	Normalized inter-class variance	RNN	Recurrent neural network
MLP	Multilayer Perceptron	SoC	Security Operations Center
STFT	Short-Time Fourier Transform	SSH	Secure Shell
SVM	Support Vector Machine	EMI	Electromagnetic Interference
LSTM	Long short-term memory	VM	Virtual Machine
CNN	Convolutional Neural Network	TCP	Transmission Control Protocol
LCD	liquid-crystal display	UDP	User Datagram Protocol
UPX	Ultimate Packer for eXecutables	AES	Advanced Encryption Standard
DES	Data Encryption Standard		

ods suffer from limited reproducibility across microarchitectures and can be influenced by software-level manipulation. As highlighted by Chenet et al. Chenet et al. (2024) in their comprehensive survey, many hardware-based malware detection approaches face common limitations, including high equipment costs, restricted datasets, and insufficient evaluation under realistic execution conditions. Recent developments in the cybercrime economy further highlight the limitations of software-based malware detection Patsakis et al. (2024). The growth of Malware-as-a-Service platforms has turned malware creation into an accessible and scalable industry in which evasion capabilities, such as code mutation, multi-layered packing, and sandbox awareness, are integrated directly into purchasable toolkits. As these services automate the generation of highly variable and analysis-resistant samples, software-level monitoring becomes increasingly unreliable. This evolution reinforces the relevance of hardware-oriented approaches, such as EM side-channel analysis, whose physical leakage cannot be concealed, intercepted, or suppressed by the malware itself.

In summary, existing literature establishes EM side-channel analysis as a viable technique for malware detection. However, three key gaps remain: (i) most studies rely on expensive laboratory equipment, (ii) Intel desktop processors remain underexplored compared to embedded platforms, and (iii) explicit detection of obfuscation techniques such as packing and virtualization has received limited attention in realistic environments. This work addresses these gaps by providing a focused, reproducible evaluation of EM-based obfuscation detection on Intel architectures using both low-cost and professional acquisition setups.

3. Setups: *EM-SENSE* vs Picoscope

To evaluate the feasibility of electromagnetic (EM) side-channel analysis for malware obfuscation detection under different cost and fidelity constraints, we employ two distinct acquisition setups: *EM-SENSE*, a low-cost and accessible prototype, and a PicoScope 5000 series professional oscilloscope-based measurement chain. Both setups are designed to capture near-field EM emissions produced by the CPU during software execution, but they differ significantly in terms of sampling capabilities, signal conditioning, cost, and deployment complexity.

The use of two contrasting acquisition systems allows us to directly assess how hardware fidelity influences signal preprocessing requirements, feature extraction, and ultimately classification performance when detecting packing and virtualization techniques.

3.1. EM-SENSE: Low-Cost Electromagnetic Acquisition

The *EM-SENSE* prototype was designed with the primary goals of accessibility, reproducibility, and low cost, enabling EM-based malware detection without the need for specialized laboratory equipment. The device is built around an Arduino UNO R4 Wifi-based acquisition platform Arduino (2023) and integrates the following components:

- An SD card module AZDelivery (2023) for real-time data storage.
- A 16x2 LCD display SunFounder (2022) for immediate signal visualization.
- A near-field probe He et al. (2016) positioned close to the CPU to capture processor emissions with minimal noise, as suggested in prior placement studies Deutschmann et al. (2005).
- A small signal amplifier and a 1 MΩ voltage divider to improve low-signal sensitivity.

All components are assembled on a modular protoboard, allowing rapid reconfiguration and easy replication without soldering. The total cost of the *EM-SENSE* setup is approximately 100€, making it suitable for reproducible research, educational environments, and exploratory deployments.

In GitHub [1] a more complete description of the elements used and their characteristics can be found.

Figure 1 shows an image of the complete setup of *EM-Sense*

[1] `https://github.com/Sergiolopezflores/EM-Side-Channels-Malware-Detection-on-Intel.git`

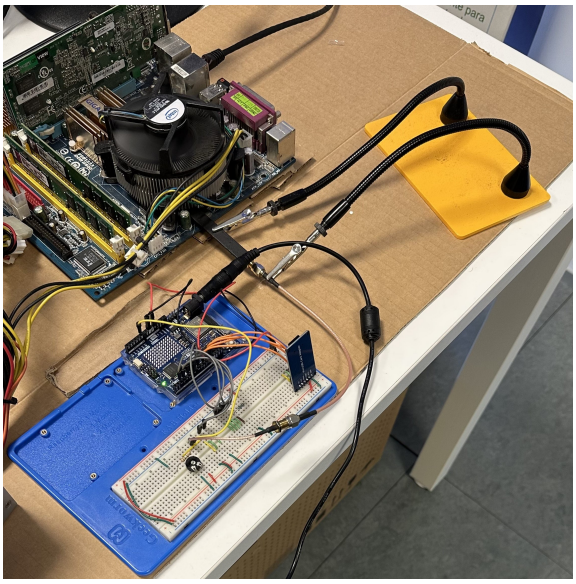


Figure 1: *EM-Sense* setup overview with component layout

3.2. Picoscope Setup

As a high-fidelity reference system, we employ a PicoScope 5000 series oscilloscope combined with professional-grade near-field probes and amplification hardware. Specifically, the setup consists of:

- A PicoScope 5443D oscilloscope connected to a host workstation for high-resolution streaming and storage.
- A Langer PA-303 +30dB amplifier, to enhance weak EM signals.
- A Langer RF-B 0.3–3 H-field probe, to ensures precise near-field measurements over a broad frequency range.

This professional configuration enables high sampling rates, wide bandwidth acquisition, and precise temporal alignment, allowing the capture of subtle spectral variations in EM emissions generated during malware execution. The total cost of this setup is approximately 6800€, excluding auxiliary laboratory infrastructure.

Figure 2 shows an image of the complete setup of *Picoscope*

Both acquisition setups target the same physical phenomenon, CPU-generated electromagnetic leakage, but represent opposite ends of the cost–fidelity spectrum. *EM-SENSE* emphasizes portability, affordability, and ease of replication, while the PicoScope-based setup prioritizes precision and bandwidth at the cost of accessibility.

Figure 3 illustrates the two measurement chains side by side, highlighting their architectural differences and emphasizing the trade-off between a compact, low-cost embedded acquisition platform and a laboratory-grade instrumentation stack delivering higher signal fidelity.

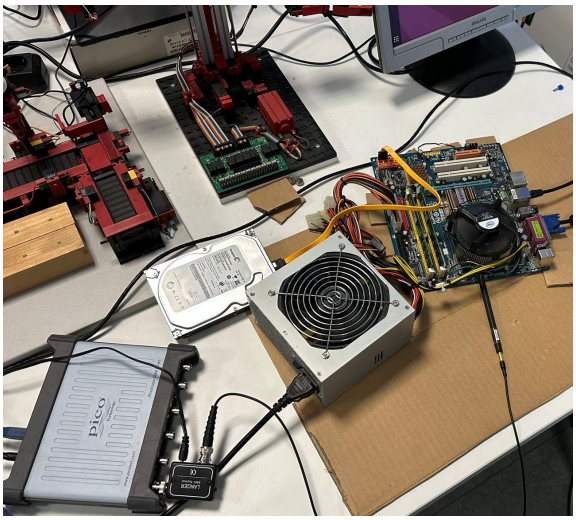


Figure 2: PicoScope setup overview with component layout

This comparative design enables a realistic evaluation of whether low-cost EM acquisition is sufficient for detecting modern malware obfuscation techniques, and how preprocessing and feature extraction strategies must adapt to hardware constraints. The implications of these differences are analyzed in detail in Sections 5 and 6.

4. EM-Powered Malware Analysis Framework

This section describes the complete malware detection framework based on electromagnetic (EM) side-channel emissions. While the overall objective is shared across both acquisition systems—detecting malware obfuscation techniques from EM leakage, the signal preprocessing and feature representation pipelines differ substantially depending on the acquisition hardware.

Specifically, the PicoScope-based setup follows a classical side-channel analysis (SCA) workflow based on time–frequency analysis and statistical feature selection, whereas *EM-SENSE* adopts a simplified, data-driven representation designed to accommodate its lower sampling rate and limited signal fidelity. Figure 4 summarizes the overall framework and highlights the two distinct preprocessing paths.

4.1. Threat Model

We consider a realistic malware analysis scenario in which analysts collect samples from open repositories, incident response feeds, or MaaS ecosystems. These samples frequently employ obfuscation techniques such as packing and code virtualization, and may include anti-debugging, anti-instrumentation, and sandbox-detection mechanisms.

We assume that analysts have physical access to the target system and can deploy external EM sensing equipment, but cannot rely on any software-level instrumentation inside the monitored host. Malware is therefore free to execute with full privileges and without restrictions. This model reflects real-world constraints and motivates the use of EM

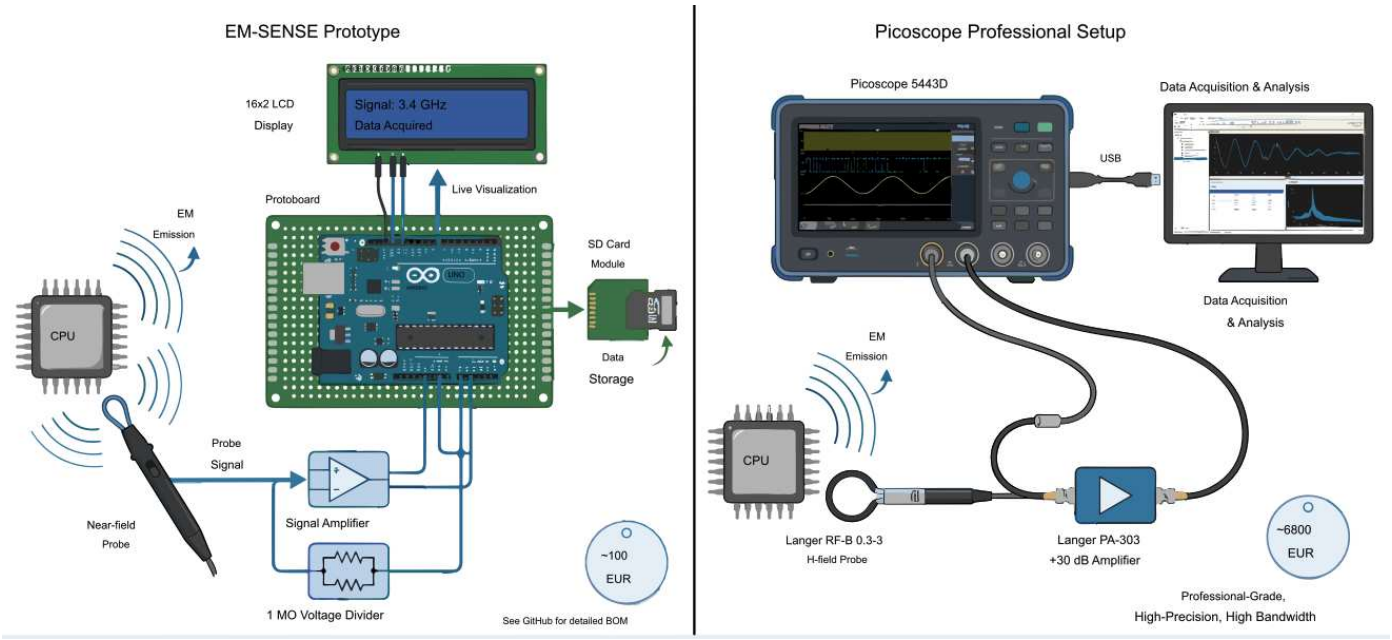


Figure 3: Comparison between the low-cost *EM-SENSE* prototype and a professional PicoScope-based instrumentation chain for near-field electromagnetic (EM) acquisition from CPUs during software execution. The *EM-SENSE* setup integrates a near-field probe, analog front-end conditioning, an Arduino-based acquisition/control unit, a 16×2 LCD for live feedback, and SD-card logging (approx. 100 EUR). The professional setup uses a Langer RF-B 0.3–3 GHz H-field probe, a Langer PA-303 (+30 dB) amplifier, and a PicoScope 5443D connected via USB to a workstation for high-fidelity acquisition and analysis (approx. 6800 EUR).

side-channel analysis, which remains observable regardless of software-level evasion techniques.

4.2. EM Signal Acquisition

During execution, malware samples and benign programs are run on bare-metal Intel hardware while EM emissions are captured using either the PicoScope-based setup or the *EM-SENSE* prototype. In both cases, a near-field probe is positioned close to the CPU package to maximize sensitivity to processor-level activity.

Although both setups observe the same physical phenomenon, they differ significantly in sampling rate, bandwidth, and signal-to-noise ratio (SNR). These differences directly influence the design of the preprocessing and feature extraction stages described below.

4.3. Preprocessing and Feature Extraction

4.3.1. PicoScope Pipeline: STFT + NICV-Based Bandwidth Selection

For the PicoScope-based setup, raw EM traces are first processed using a Short-Time Fourier Transform (STFT) to obtain a time–frequency representation of the signal. This step exposes spectral components associated with program execution phases, such as memory access patterns, control-flow activity, and cryptographic operations.

Given the high sampling rate and wide bandwidth provided by the PicoScope, the resulting spectrogram contains a large number of frequency bins, many of which are redundant or dominated by noise. To isolate the most discriminative spectral components, we apply Normalized Inter-Class Variance (NICV), a statistical metric commonly used in side-channel analysis to quantify class separability.

NICV is computed for each frequency bin across classes (e.g., packed vs. unpacked, virtualized vs. non-virtualized), and frequency bands exhibiting the highest inter-class variance are selected. This process effectively performs bandwidth selection, retaining only those spectral regions that maximize discrimination while reducing dimensionality.

The final feature representation for the PicoScope pipeline consists of a fixed-length vector formed by aggregating normalized spectral energy values from the selected frequency bands. This representation is well suited to both classical machine learning and deep learning classifiers and preserves physically meaningful EM characteristics.

4.3.2. EM-SENSE Pipeline: Direct One-Hot Encoded Representation

In contrast, the *EM-SENSE* prototype operates at a significantly lower sampling rate and exhibits higher susceptibility to environmental noise and electromagnetic interference. Under these constraints, applying fine-grained frequency-domain analysis and statistical bandwidth selection becomes less effective and computationally impractical.

Instead, *EM-SENSE* adopts a direct data-driven representation. Raw time-domain EM traces are segmented, normalized, and transformed into a fixed-length one-hot encoded format, where each trace is represented as a high-dimensional binary or discretized vector. This representation does not attempt to explicitly isolate frequency bands, but instead allows the learning model to implicitly capture

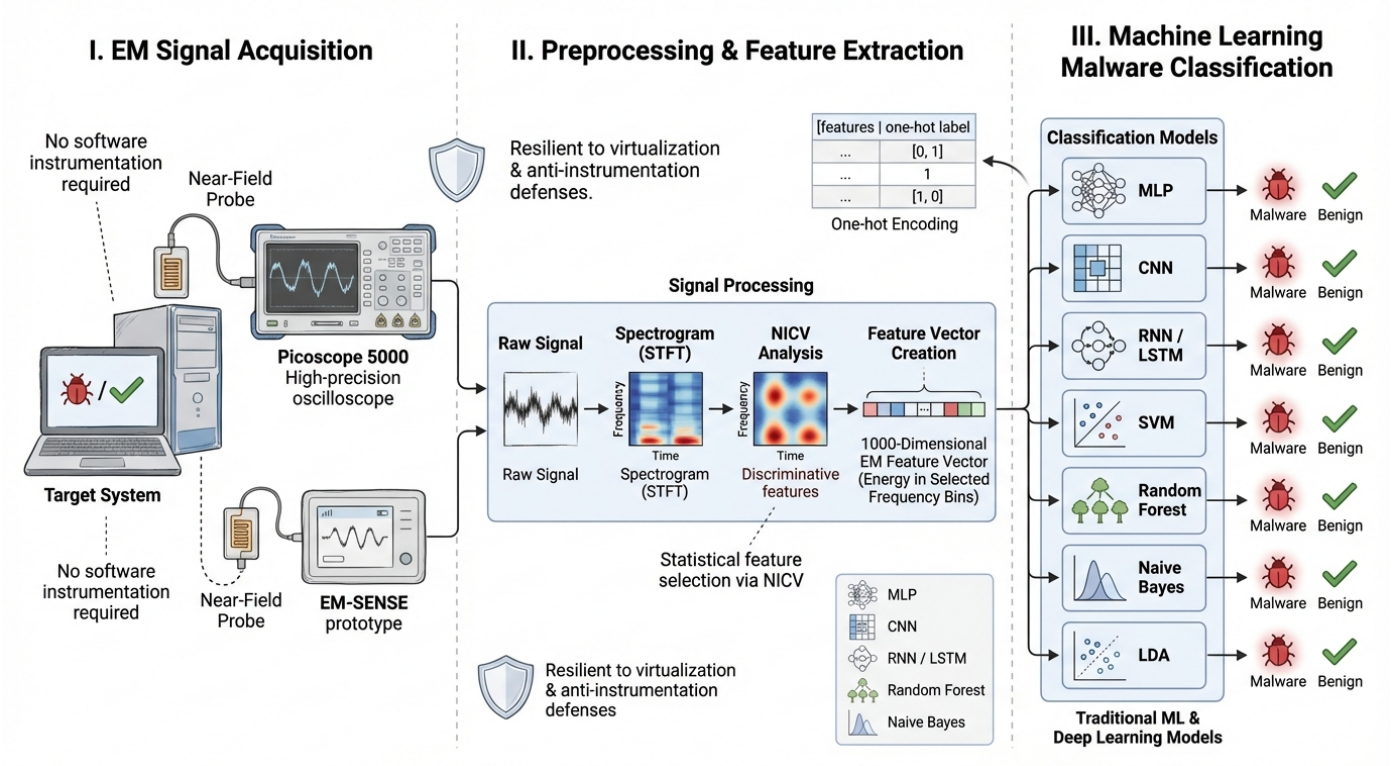


Figure 4: End-to-end pipeline for non-intrusive malware detection from near-field electromagnetic (EM) emanations. **(I) EM signal acquisition:** a near-field probe captures emissions from the target system during program execution without requiring software instrumentation, using either a high-precision PicoScope 5000-series oscilloscope or the low-cost *EM-SENSE* prototype. **(II) Preprocessing and feature extraction:** raw traces are transformed into time–frequency representations (STFT spectrograms), followed by NICV-based analysis to identify discriminative frequency regions and produce a fixed-length EM feature vector (e.g., energy in selected frequency bins, on the order of 10^3 dimensions). Labels are encoded (e.g., one-hot) for supervised learning. **(III) Machine-learning classification:** the resulting feature vectors are fed to traditional and deep learning classifiers (MLP, CNN, RNN/LSTM, SVM, Random Forest, Naive Bayes, LDA) to discriminate benign from malware execution.

discriminative patterns from the encoded signal structure.

Although *EM-SENSE* does not perform explicit spectral decomposition, differences between packed and non-packed executions are already observable at the raw signal level. Figures 5 and 6 illustrates representative raw EM traces captured during the execution of packed and non-packed binaries using *EM-SENSE*. Even under lower sampling resolution, visible amplitude variations and structural differences can be observed, particularly during unpacking phases.

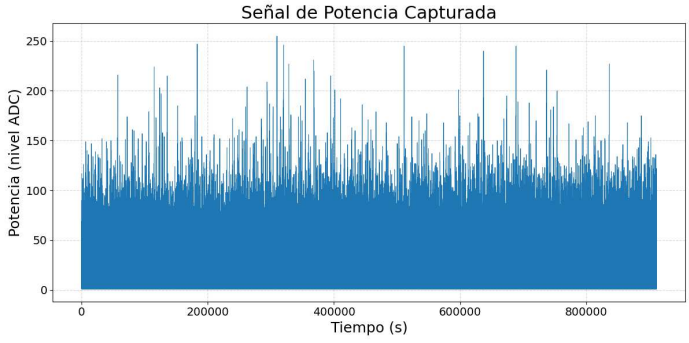


Figure 5: Raw EM signal for packed using *EM-SENSE*

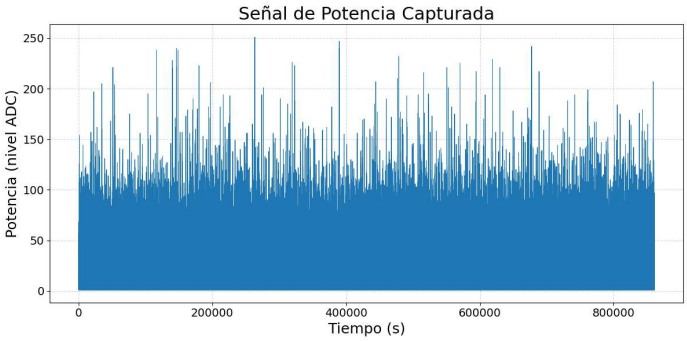


Figure 6: Raw EM signal for not packed using *EM-SENSE*

These visual differences motivate the use of direct temporal encoding, as discriminative information is embedded in signal amplitude modulation and temporal structure rather than in isolated narrow frequency bands. The one-hot encoded representation therefore provides a simplified yet effective mechanism to capture these patterns without requiring computationally expensive spectral transforms.

This approach trades interpretability and explicit frequency selectivity for simplicity and robustness under low-fidelity acquisition. Importantly, this representation is not

used for the PicoScope pipeline, as the availability of high-resolution spectral information makes NICV-based feature selection both feasible and preferable in that case.

4.4. Feature Representation Summary

The two preprocessing pipelines can be summarized as follows:

- *PicoScope setup:* Raw EM traces → STFT spectrogram → NICV-based bandwidth selection → fixed-length spectral feature vector
- *EM-SENSE setup:* Raw EM traces → normalization and segmentation → one-hot encoded vector representation

Although the resulting feature spaces differ in structure and dimensionality, both are designed to preserve execution-dependent EM characteristics relevant to detecting malware obfuscation techniques.

4.5. Classification Stage

The extracted feature representations are used as input to supervised classification models tasked with binary detection of packing and virtualization. Both classical machine learning models (e.g., Naïve Bayes, Support Vector Machines, Random Forests) and deep learning architectures (e.g., MLPs, CNNs, and RNN-based models) are evaluated.

The objective of this work is not to propose novel learning architectures, but to assess how acquisition fidelity and preprocessing strategy impact classification performance. In particular, we analyze whether low-cost acquisition combined with simplified feature representations can achieve detection performance comparable to that obtained with high-end instrumentation and SCA-inspired preprocessing.

5. Experiments

This section describes the experimental environment used to evaluate electromagnetic (EM) side-channel–based detection of malware obfuscation techniques. We detail the target platform, malware samples, benign workloads, data acquisition procedure, and dataset construction for both acquisition setups: the professional PicoScope-based system and the low-cost *EM-SENSE* prototype.

The experimental design aims to balance realism, reproducibility, and isolation, ensuring that captured EM traces reflect genuine execution behavior while remaining safe and controllable.

5.1. Target Platform

All experiments were conducted on a physical (bare-metal) desktop computer equipped with an Intel processor, representative of general-purpose systems commonly targeted by modern malware. The system ran a 64-bit Linux operating system, equipped with an Intel Core 2 Quad CPU @ 2.4 GHz × 4, running Ubuntu 22.04.5 LTS and executed over one hundred background processes and services during experiments, including system daemons and user-level applications, to reproduce realistic multi-process conditions. This configuration was selected for its stability, accessibility, and representativeness of general-purpose computing systems that are frequently targeted by malware.

Virtualization and sandboxing technologies were deliberately avoided. This design choice prevents virtualization artifacts and eliminates common triggers used by malware to alter behavior under analysis environments. The target machine was isolated from external networks and remotely accessed to avoid physical interference during EM measurements.

Near-field EM probes were mounted on a fixed mechanical support to ensure consistent placement relative to the CPU package throughout all experiments.

5.2. Malware Categories

The malware dataset used in this study is constructed from three well-known and widely studied Linux malware families: Mirai, Bashlite, and Gonnacry. These families were selected due to their prevalence, behavioral diversity, and continued relevance in modern threat landscapes, as well as their frequent appearance in empirical malware research Antonakakis et al. (2017); Böck et al. (2023); Zhu et al. (2022).

Rather than treating malware families as classification targets, they are used here as base executables from which multiple obfuscation variants are derived. This design choice allows the experimental analysis to focus specifically on the impact of obfuscation techniques, while controlling for underlying malware functionality.

5.2.1. Base Malware Samples

Mirai is a botnet malware family responsible for some of the most disruptive distributed denial-of-service (DDoS) attacks observed to date. It is characterized by aggressive network scanning, credential brute-forcing, and persistent command-and-control communication. Its large-scale impact and continuous evolution have been extensively documented in the literature Antonakakis et al. (2017); Böck et al. (2023); Zhu et al. (2022).

Bashlite is a lightweight Linux botnet malware family closely related to Mirai, featuring network flooding routines and simple but persistent execution loops. Despite its relative simplicity, Bashlite remains prevalent in IoT and Linux ecosystems and is frequently used in experimental evaluations of malware behavior.

Gonnacry is an open-source ransomware sample designed to encrypt user files using symmetric cryptography. The variants of Gonnacry employ different cryptographic algorithms, introducing intensive computation phases that differ fundamentally from network-centric botnet malware. The inclusion of ransomware reflects the increasing targeting of Linux-based systems by modern ransomware campaigns Korac et al. (2024).

Together, these three families provide a diverse execution profile, covering both network-intensive behavior (Mirai, Bashlite) and computation-heavy behavior (Gonnacry), while remaining representative of contemporary Linux malware ecosystems.

5.2.2. Obfuscation Variants

For each base malware sample, multiple obfuscation variants were generated to support the two binary classification tasks evaluated in this work. First, unobfuscated binaries were executed as baseline references. For the packing scenario, original malware binaries were transformed using common executable packing techniques; in particular, the UPX packer was applied with standard compression and executable protection options. Each packed binary was paired with its corresponding unpacked version, ensuring that observed differences in electromagnetic (EM) traces could be attributed to the packing process rather than to changes in functional behavior. For the virtualization scenario, malware binaries protected using code virtualization techniques were executed alongside their non-virtualized counterparts. Virtualized binaries are based on custom bytecode interpreters and virtual execution loops, introducing distinctive execution patterns at the CPU level. These structural differences are expected to manifest in EM side-channel emissions, even when software-level behavior is intentionally obscured.

Each obfuscated sample is paired with its corresponding unobfuscated version, ensuring that observed differences in electromagnetic (EM) traces can be attributed primarily to the obfuscation technique rather than to functional differences in malware behavior. This pairing strategy follows best practices in malware analysis and mitigates confounding effects caused by behavioral diversity.

5.2.3. Dataset Rationale

The choice of Mirai, Bashlite, and Gonnacry is motivated by their prevalence, functional diversity, and relevance to current threat models. Botnet malware such as Mirai and Bashlite continues to dominate Linux and IoT threat ecosystems, while ransomware increasingly targets Linux-based servers and cloud infrastructures Antonakakis et al. (2017); Korac et al. (2024). Moreover, these families are commonly distributed through Malware-as-a-Service platforms, where automated packing and virtualization are routinely applied to avoid detection Patsakis et al. (2024).

By deriving obfuscated and unobfuscated variants from the same base malware samples, the dataset isolates the effects of packing and virtualization, making it well suited for evaluating electromagnetic side-channel–based detection of modern malware evasion techniques.

5.3. Benign Workloads

To prevent the learning problem from collapsing into an *idle vs. active* discriminator, malware executions were collected under concurrent benign activity. The goal is not to model a specific user profile, but to introduce contention and variability from common CPU and I/O sources so that EM traces reflect a multi-process host state.

Benign activity combines short-lived utilities and longer-running workloads. The utilities cover file-system actions and text/system inspection (e.g., `ls`, `find`, `grep`, `dmesg`, `objdump`), while additional programs induce network activity (`wget`, `ss`, `ip`), compression and archiving (`bzip2`, `tar`, `gunzip`), bulk I/O (`dd`), scripting workloads (`python`), and multimedia-oriented execution with sustained runtime (e.g., audio playback and video muxing when available). This mix spans short to long runtimes and low to high CPU utilization, and it generates background interference patterns that overlap with phases commonly present in malware execution.

Several hardware-assisted malware/intrusion detection studies-including EM-based approaches-evaluate under comparatively controlled benign conditions, often relying on a limited set of benign programs or fixed benchmarks Chawla et al. (2021); Khan et al. (2019a); Nazari et al. (2017); Sehatbakhsh et al. (2019). Similar constraints also appear in other hardware-signal modalities such as performance counters Singh et al. (2017). Our benign construction is therefore intended to reduce reliance on collection-time artifacts and to better reflect concurrent activity during deployment.

5.4. Data Acquisition Procedure

For each experiment, samples were repeatedly executed while benign activity was interleaved according to a randomized run order. Execution start times were randomized to avoid systematic alignment across traces and to reduce dependence on transient system state. Probe placement and the mechanical mounting were kept fixed throughout the campaign to avoid geometry-induced variability.

Electromagnetic emissions were acquired using one of the setups described in Section 3:

- *PicoScope setup:* Each execution produced a dedicated high-bandwidth trace recorded at high sampling rate and stored for offline processing. This acquisition preserves wide spectral content suitable for time-frequency analysis and statistical feature selection.

- *EM-SENSE setup:* Each execution produced a lower-rate time-domain trace logged locally to an SD card. Acquisition was externally triggered and recorded passively, reflecting the constraints of an embedded, low-cost measurement path without continuous synchronization.

5.5. Dataset Construction and Splitting

Captured traces were processed using the pipeline associated with the acquisition hardware (Section 4), as bandwidth and signal-to-noise characteristics differ substantially between setups:

- *PicoScope dataset:* Raw traces were converted into time-frequency representations using STFT. Normalized Inter-Class Variance (NICV) was then computed across frequency bins to select discriminative spectral regions. Final inputs are fixed-length vectors of normalized spectral energy over the selected bands.
- *EM-SENSE dataset:* Raw traces were segmented and normalized, then mapped to a fixed-length, one-hot encoded representation. This avoids explicit spectral decomposition and provides a consistent input format under the prototype's sampling constraints.

For each task, the traces were labeled by obfuscation condition (packed vs. unpacked; virtualized vs. non-virtualized) and evaluated independently. Data were split into 80% training and 20% validation at the execution-instance level, ensuring that no trace from the same run appears in both partitions. The same splitting protocol was applied across setups to support a like-for-like comparison.

5.6. Experimental Scope

The evaluation targets binary detection of two obfuscation mechanisms–packing and code virtualization–rather than malware family attribution. This isolates the effect of obfuscation on EM leakage and matches the study question: whether external EM sensing provides separable structure for these transformations under a non-instrumented threat model.

By evaluating a laboratory-grade chain and the low-cost *EM-SENSE* prototype under the same execution protocol and split strategy, the experiments quantify how acquisition fidelity constrains representation choices and downstream detection performance.

6. Results Analysis

This section presents a comprehensive evaluation of electromagnetic (EM) side-channel–based detection of malware obfuscation techniques under both acquisition setups. We report results for virtualization and packing detection using deep learning (DL) and classical machine learning (ML) approaches within a unified framework. All experiments were conducted using the preprocessing pipelines described in Section 4, and models were trained and validated on a controlled virtualized server environment to ensure consistent computational conditions and reproducibility. Accuracy is reported for each model, and representative confusion matrices are analyzed to provide deeper insight into classification behavior.

6.1. PicoScope-Based Detection Performance

Table 2 summarizes the classification accuracy obtained with PicoScope-acquired EM traces.

6.1.1. Virtualization Detection

For the virtualization detection task, deep learning models (CNN, MLP, RNN) achieve accuracies between 67% and 71%. Although convolutional and recurrent architectures are able to exploit the temporal–spectral structure in STFT representations, their performance remains significantly below that of classical models.

LDA-based classical approaches clearly dominate this scenario. In particular, LDA+SVM achieves 94.02% accuracy, followed by LDA+NB at 92.43%. The strong performance of these models indicates that virtualization introduces highly discriminative execution patterns that are well captured by NICV-selected frequency bands. Since virtualization relies on persistent virtual machine loops and dispatcher routines, the resulting EM emissions exhibit stable and repetitive structures that are largely linearly separable after dimensionality reduction.

6.1.2. Packing Detection

Packing detection under PicoScope acquisition achieves the highest overall performance. Deep learning models reach approximately 91% accuracy, indicating that high-resolution spectral information significantly enhances feature richness.

However, classical models again outperform DL. LDA +NB achieves 99.20% accuracy, LDA+RF reaches 98.76%, and LDA+SVM achieves 98.31%. The confusion matrix corresponding to the PicoScope packing detection with SVM model scenario (Figure 7a) exhibits an almost perfectly diagonal structure. A total of 255 original samples are correctly classified with no false positives, while 241 packed samples are correctly detected, with only three misclassifications. This near-perfect separation confirms that unpacking routines generate strong and consistent spectral signatures when captured at high bandwidth.

The minimal classification error suggests that the NICV-based bandwidth selection effectively isolates frequency components associated with decompression loops, memory copying, and execution transfer phases. These artifacts are particularly visible in high-fidelity measurements, explaining the marginal but consistent improvement over *EM-SENSE*.

6.2. EM-SENSE-Based Detection Performance

Despite its substantially lower sampling rate and simplified preprocessing pipeline, *EM-SENSE* achieves competitive performance, especially for virtualization detection, Table 2.

6.2.1. Virtualization Detection

Deep learning models trained on one-hot encoded time-domain traces achieve moderate performance, with accuracies ranging from 58% to 64%. These results suggest that end-to-end feature extraction from raw temporal encodings is insufficient to fully capture discriminative patterns under low-fidelity acquisition.

Scenarios	#	*CNN*	*MLP*	*RNN*	*LDA+NB*	*LDA+SVM*	*LDA+RF*
Virtualization	2	**70.56**	67.12	69.47	92.43	**94.02**	77.48
Packer	2	**90.86**	90.56	90.62	**99.20**	98.31	98.76

Table 2: Accuracy (AC) for CNN, MLP, RNN from DL and NB, SVM, and RF (after LDA) from ML. Best accuracy per scenario in bold. "#" indicates number of classes.

Scenarios	#	*CNN*	*MLP*	*RNN*	*LDA+NB*	*LDA+SVM*	*LDA+RF*
Virtualization	2	**64.35**	60.87	58.21	**90.32**	89.85	85.47
Packer	2	**62.32**	60.22	59.43	89.23	90.47	**95.68**

Table 3: Accuracy (AC) for CNN, MLP, RNN from DL and NB, SVM, and RF (after LDA) from ML. Best accuracy per scenario in bold. "#" indicates number of classes.

In contrast, classical LDA-based models significantly outperform deep architectures. Both LDA+NB and LDA +SVM reach 90% accuracy, demonstrating that virtualization introduces strong execution-level artifacts that remain detectable even without explicit spectral analysis.

The confusion matrix corresponding to the *EM-SENSE* virtualization detection using LDA+NB (Figure 7b) provides deeper insight into classifier behavior. The model correctly classifies 161 original samples and 137 virtualized samples, while misclassifying 14 original samples as virtualized and 19 virtualized samples as original. Although a limited degree of overlap exists between classes, the matrix remains strongly diagonal, confirming reliable separability. The presence of both false positives and false negatives suggests that certain virtualization-induced execution segments may temporally overlap with benign instruction patterns under lower sampling resolution. Nevertheless, overall classification stability remains high, with balanced precision and recall across both classes.

These results confirm that virtualization detection is feasible under constrained acquisition conditions, although slightly more sensitive to signal degradation than in the high-fidelity scenario.

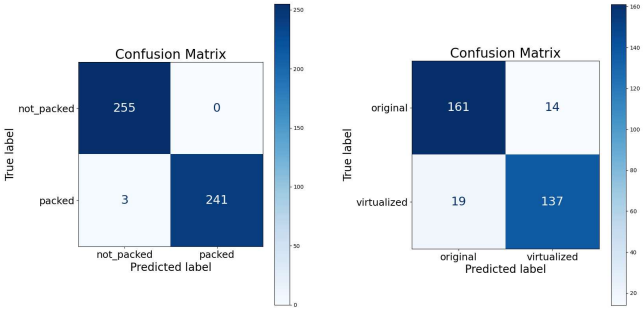
6.2.2. Packing Detection

For packing detection, *EM-SENSE* achieves a maximum accuracy of 96% using LDA+RF, with LDA+NB and LDA+SVM reaching approximately 89–90%. Deep learning models remain near 60% accuracy, highlighting the difficulty of extracting discriminative patterns from raw time-domain encodings when unpacking phases are short-lived.

Compared to PicoScope results, packing detection shows a modest performance reduction. This gap can be attributed to reduced bandwidth and sampling resolution, which may partially obscure transient decompression phases. Nevertheless, the high performance of classical models confirms that even low-cost EM acquisition preserves sufficient execution-level information for reliable discrimination.

6.3. Cross-Setup Comparative Analysis

A direct comparison between PicoScope and *EM-SENSE* reveals three key observations.



(a) SVM packed confusion matrix using PicoScope. (b) NB virtualized confusion matrix using *EM-SENSE*.

Figure 7: confusion matrices.

First, virtualization detection remains robust across both setups, with only a moderate decrease in accuracy (94.02% vs. 90%). This indicates that virtualization artifacts are strong, repetitive, and structurally stable, making them resilient to acquisition fidelity limitations.

Second, packing detection benefits more significantly from high-resolution acquisition. PicoScope achieves up to 99.20% accuracy, while *EM-SENSE* reaches 96% at best. The improvement suggests that unpacking phases are more frequency-dependent and may be partially smoothed in lower-bandwidth recordings.

Third, classical models consistently outperform deep learning in most scenarios. This behavior suggests that, in binary obfuscation detection tasks, discriminative information is concentrated in a limited feature space that does not require high-capacity nonlinear modeling. The combination of STFT and NICV in the PicoScope pipeline further enhances this separability.

6.4. Deep Learning vs. Classical Modeling

Across both acquisition setups, classical LDA-based models consistently demonstrate superior or at least comparable performance relative to deep neural networks. This trend can be attributed to several factors. First, the binary nature of the classification tasks reduces the necessity for high-capacity hierarchical feature representations, as the

decision boundary between classes is relatively well defined. Second, NICV-based bandwidth selection effectively concentrates discriminative information into a low-dimensional feature space, allowing linear and ensemble-based classifiers to operate on highly informative projections. Third, the dataset size remains moderate in comparison to the complexity of deep learning architectures, which limits the potential generalization gains that might otherwise justify deeper models. Additionally, obfuscation techniques such as packing and virtualization introduce structured and repetitive execution patterns that are naturally well captured through linear projections after appropriate feature selection. Deep learning models show improved performance primarily in the PicoScope packing scenario, where richer spectral information enables convolutional filters to exploit localized time–frequency patterns. However, even in this case, classical models remain competitive while requiring significantly lower computational cost and training overhead.

6.5. Exploratory Evaluation of Embedded Platforms for Future EM-SENSE Deployment

In addition to the core experimental validation presented in the previous sections, we conducted an exploratory evaluation of different hardware platforms to assess the feasibility of deploying the trained models in a future evolution of the *EM-SENSE* system. This analysis is not intended to extend the experimental claims of EM-based obfuscation detection, but rather to evaluate practical deployment considerations should *EM-SENSE* transition toward autonomous, on-device inference.

Table 4 reports, for each platform and binary task (Packer and Virtualized), the best-performing classical model and the best-performing deep learning model, along with their respective per-epoch training times. Since both tasks involve only two classes (packed vs. not packed, virtualized vs. not virtualized), performance differences across platforms are primarily computational rather than accuracy-driven.

Across all platforms, classical models, specifically LDA +NB for the Packer task and LDA+SVM for the Virtualized task, consistently achieve the highest accuracy. Notably, accuracy remains stable across hardware environments: 0.99 for Packer and 0.94–0.99 for Virtualized, depending on the platform. This invariance confirms that model generalization is unaffected by execution hardware and that inference portability is feasible without retraining or architectural modification.

Training time, however, varies significantly. On the virtualized server, classical models complete training in approximately one second per epoch. On embedded devices, per-epoch time increases proportionally to computational capacity: three seconds on the Raspberry Pi 5, nine seconds on the Jetson Nano, and five to six seconds on the Jetson Nano Orin. Despite these increases, execution remains lightweight and well within acceptable limits for offline training or periodic retraining scenarios.

Deep learning models exhibit a different behavior. While CNN-based models achieve stable accuracy for Packer detection (0.91 across all platforms), their performance for Virtualized detection remains significantly lower (0.71). Unlike classical models, CNN training time varies more substantially across platforms. On the Raspberry Pi 5, training requires five seconds per epoch for Packer and three seconds for Virtualized. On the Jetson Nano, training extends to 10–11 seconds per epoch. The Jetson Nano Orin shows improved efficiency, reducing CNN training time to two seconds per epoch while maintaining identical accuracy.

Importantly, the accuracy values for both ML and DL models remain constant across platforms, confirming that hardware differences do not affect classification capability but only computational cost. This behavior reinforces the suitability of classical LDA-based models for edge deployment, as they preserve maximal accuracy while requiring minimal processing resources.

From a design perspective, these results must be interpreted in the context of the current experimental scope, which involves binary classification tasks (packed vs. not packed and virtualized vs. not virtualized). In this relatively low-complexity setting, both the Raspberry Pi 5 and GPU-enabled platforms such as the Jetson Nano Orin represent viable deployment options, as classification accuracy remains stable across devices and computational requirements are modest. Classical LDA-based models, in particular, execute efficiently on both architectures without any degradation in performance. However, if future iterations of *EM-SENSE* were extended to more complex multi-class scenarios, the computational burden, especially for deep learning architectures, would likely increase, making GPU acceleration more relevant. In such cases, Jetson-class devices would provide a clearer advantage. For the binary obfuscation detection tasks considered in this work, however, both platforms offer an adequate balance between performance and deployability.

It is important to emphasize that this embedded-platform evaluation does not modify the primary conclusions of this work. The core contribution remains the demonstration that EM side-channel emissions enable reliable detection of malware obfuscation techniques under both professional and low-cost acquisition setups. The present analysis merely indicates that, should future iterations of *EM-SENSE* aim for autonomy, the computational requirements for on-device inference are modest and technically feasible.

7. Discussion

This work investigates the feasibility of detecting malware obfuscation techniques, specifically packing and code virtualization, through electromagnetic (EM) side-channel analysis under both high-fidelity and low-cost acquisition setups. The results provide several insights regarding de-

Table 4: **Best model per platform and task.** For each hardware platform and scenario, we report the model with the highest accuracy and its per-epoch training time (seconds). Ties are resolved by selecting the configuration with the lowest per-epoch time (deployability criterion).

Platform	Task	Best model(ML/DL)	Accuracy	Time/epoch (s)
VM / Virtualized Server	Packer	LDA+NB	0.99	1
		CNN	0.91	1
	Virtualized	LDA+SVM	0.94	1
		CNN	0.71	1
Raspberry Pi 5	Packer	LDA+NB	0.99	3
		CNN	0.91	5
	Virtualized	LDA+SVM	0.94	2
		CNN	0.71	3
Jetson Nano	Packer	LDA+NB	0.99	9
		CNN	0.91	10
	Virtualized	LDA+SVM	0.94	9
		CNN	0.71	11
Jetson Nano Orin	Packer	LDA+NB	0.99	6
		CNN	0.91	2
	Virtualized	LDA+SVM	0.94	5
		CNN	0.71	2

tectability, acquisition constraints, modeling choices, and practical deployment considerations.

7.1. Robustness of EM Leakage Against Software Obfuscation

A central finding of this study is that software-level obfuscation techniques cannot conceal the physical side effects produced by processor activity. Both packing and virtualization introduce execution patterns that manifest in EM emissions, even when static and dynamic analysis techniques are significantly hindered.

Virtualization, in particular, generates highly stable and repetitive execution structures due to the presence of dispatcher loops and instruction handlers. These repetitive patterns result in consistent EM signatures that remain detectable even under reduced sampling resolution, as demonstrated by the strong performance achieved with *EM-SENSE*. Packing, while more transient in nature, still introduces computational bursts during decompression and memory transfer phases that become highly distinguishable under high-bandwidth acquisition.

These observations reinforce the fundamental premise of hardware-assisted malware detection: physical side channels provide an observation layer that is inherently external to the software execution stack and therefore resistant to traditional evasion strategies.

7.2. Impact of Acquisition Fidelity

The comparative evaluation between PicoScope and *EM-SENSE* reveals that acquisition fidelity influences performance margins but does not fundamentally alter detectability.

High-resolution acquisition improves accuracy, particularly for packing detection, where short-lived unpacking phases benefit from wider bandwidth and improved signal-to-noise ratio. However, the performance degradation observed under *EM-SENSE* remains moderate, especially for virtualization detection. This suggests that certain obfuscation artifacts are structurally strong enough to survive reduced sampling conditions.

Importantly, this finding has practical implications: while professional-grade instrumentation maximizes classification accuracy, low-cost EM acquisition can still provide reliable detection, significantly lowering the barrier to reproducible research and real-world deployment.

7.3. Classical Models vs. Deep Learning Architectures

Across both setups, classical LDA-based models consistently outperform or match deep neural networks in the evaluated binary tasks. This behavior highlights an important methodological insight: model complexity should be aligned with feature structure rather than assumed beneficial by default.

The binary nature of the tasks, combined with NICV-based bandwidth selection, concentrates discriminative information into a low-dimensional feature space. Under these conditions, linear and ensemble-based classifiers are sufficient to achieve high accuracy while avoiding overfitting and excessive computational overhead.

Deep learning models demonstrate competitive performance primarily in the PicoScope packing scenario, where richer spectral representations allow convolutional filters to exploit localized time–frequency structures. Nevertheless, the marginal performance gains do not offset the increased computational cost in most deployment scenarios.

This observation supports the argument that, for obfuscation-focused EM detection, carefully engineered feature extraction may be more impactful than increasing model depth.

7.4. Deployment Implications and Hardware Evolution

The exploratory evaluation of alternative embedded platforms demonstrates that classification accuracy remains stable across hardware environments. For the binary tasks considered in this work, both Raspberry Pi–class devices and GPU-enabled Jetson platforms represent viable deployment options.

The computational footprint of LDA-based models is sufficiently small to enable near-real-time inference on lightweight embedded processors. This opens the possibility of evolving *EM-SENSE* from a passive acquisition prototype into a semi-autonomous or fully autonomous detection device.

However, this conclusion must be interpreted within scope. The tasks evaluated in this study involve only two classes. Should future research extend to more complex multi-class scenarios, such as family-level classification or multi-stage behavioral inference, computational requirements would likely increase. In such cases, GPU acceleration may become necessary to sustain practical training and inference times.

7.5. Limitations and Threats to Validity

While the results are encouraging, several limitations must be acknowledged.

First, the dataset focuses on three Linux malware families and two specific obfuscation techniques. Although these techniques are widely used in practice, the generalization of results to other malware families, custom packers, or advanced virtualization engines requires further validation.

Second, experiments were conducted in a controlled environment with consistent probe placement and stable hardware configuration. Real-world deployments may introduce environmental electromagnetic interference, hardware variability, or background workload fluctuations that could affect signal quality.

Third, models were trained and validated using data collected from the same hardware platform. Although care was taken to avoid data leakage between training and validation sets, cross-hardware generalization remains an open research direction.

Finally, this study evaluates offline detection. Real-time continuous monitoring introduces additional challenges related to segmentation, windowing strategy, and latency constraints that are not addressed in the current work.

7.6. Broader Implications

The findings of this work contribute to the broader discussion on hardware-assisted cybersecurity defenses. As malware continues to incorporate increasingly sophisticated obfuscation and anti-analysis mechanisms, purely software-based detection strategies face growing limitations.

EM side-channel analysis provides a complementary perspective that is fundamentally orthogonal to traditional approaches. By observing physical leakage rather than software artifacts, detection mechanisms become resilient to many forms of runtime evasion.

Moreover, demonstrating that low-cost acquisition can achieve competitive results suggests that hardware-assisted detection need not remain confined to specialized laboratory environments. This democratization of EM-based analysis could enable new defensive paradigms in edge computing, IoT security, and high-assurance systems.

8. Conclusion

This work presented a focused evaluation of electromagnetic (EM) side-channel analysis for the detection of malware obfuscation techniques on Intel desktop processors. Specifically, we investigated the feasibility of distinguishing between packed and non-packed binaries, as well as virtualized and non-virtualized executables, under two acquisition setups: a professional PicoScope-based measurement chain and the low-cost *EM-SENSE* prototype.

The results demonstrate that EM side-channel emissions preserve discriminative information even when malware employs advanced software-level obfuscation. Virtualization detection achieved up to 94% accuracy under high-fidelity acquisition and remained robust at 90% under low-cost acquisition. Packing detection reached up to 99.20% accuracy with PicoScope measurements and 96% with *EM-SENSE*. These findings confirm that obfuscation techniques cannot conceal the physical execution patterns they induce at the hardware level.

A key outcome of this study is that classical LDA-based machine learning models consistently outperform or match deep learning architectures for the evaluated binary tasks. This highlights the importance of feature engineering and discriminative bandwidth selection over model complexity in EM-based obfuscation detection. Furthermore, the strong performance of lightweight classifiers suggests that near-real-time deployment is feasible without requiring high-end computational infrastructure.

The comparative analysis between PicoScope and *EM-SENSE* shows that while professional-grade instrumentation provides marginal accuracy improvements, particularly for packing detection, low-cost acquisition remains highly effective, especially for virtualization detection. This significantly lowers the barrier for reproducible research and practical deployment of EM-assisted malware detection systems.

An exploratory evaluation of alternative embedded platforms further indicates that, for binary obfuscation detection, both Raspberry Pi–class devices and GPU-enabled Jetson platforms represent viable deployment options. Although more complex multi-class tasks may require hardware acceleration, the current results demonstrate that

autonomous EM-based detection devices are technically feasible.

Overall, this work reinforces the potential of electromagnetic side-channel analysis as a resilient and complementary defense mechanism against modern malware evasion techniques. By demonstrating robust detection under both professional and low-cost acquisition setups, we contribute to bridging the gap between laboratory-grade hardware security research and practical, deployable cybersecurity solutions.

Future work will focus on expanding the dataset to additional malware families and obfuscation engines, evaluating cross-hardware generalization, and investigating continuous real-time monitoring strategies. These directions aim to further assess the scalability and robustness of EM-based detection in real-world operational environments.

Acknowledgements

This publication is part of the project "CiberIA: Investigación e Innovación para la Integración de Ciberseguridad e Inteligencia Artificial" (Proyecto C079/23), financed by "European Union NextGeneration-EU", the Recovery Plan, Transformation and Resilience, through "INCIBE". It has also been partially supported by the project SecAI (PID2022-139268OB-I00) funded by the "Spanish Ministerio de Ciencia e Innovacion", and "Agencia Estatal de Investigacion." Funding for open access charge: Universidad de Málaga / CBUA.

References

Antonakakis, M., April, T., Bailey, M., et al., 2017. Understanding the mirai botnet, in: USENIX Security Symposium.

Arduino, 2023. Arduino UNO R4 WiFi Datasheet. Available: https://docs.arduino.cc.

AZDelivery, 2023. AZDelivery SPI SD Card Module for Arduino. Available online.

Berger, V.W., Zhou, Y., 2014. Kolmogorov–smirnov test: Overview. Wiley statsref: Statistics reference online .

Böck, B., et al., 2023. The end of the canonical iot botnet: A measurement study of mirai's descendants. arXiv preprint arXiv:2309.01130 .

Chawla, N., Kumar, H., Mukhopadhyay, S., 2021. Machine learning in wavelet domain for electromagnetic emission based malware analysis. IEEE Transactions on Information Forensics and Security 16, 3426–3441.

Chenet, C.P., Savino, A., Di Carlo, S., 2024. A survey on hardware-based malware detection approaches. IEEE Access 12, 54115–54128.

Clark, S.S., Ransford, B., Rahmati, A., Guineau, S., Sorber, J., Xu, W., Fu, K., 2013. WattsUpDoc: Power side channels to nonintrusively discover untargeted malware on embedded medical devices, in: 2013 USENIX Workshop on Health Information Technologies (HealthTech 13).

Demme, J., Maycock, M., Schmitz, J., Tang, A., Waksman, A., Sethumadhavan, S., Stolfo, S., 2013. On the feasibility of online malware detection with performance counters. ACM SIGARCH computer architecture news 41, 559–570.

Deutschmann, B., Pitsch, H., Langer, G., 2005. Near field measurements to predict the electromagnetic emission of integrated circuits, in: Proc. 5th Int. Workshop Electromagn. Compat. Integr. Circuits, pp. 27–32.

Ding, F., Li, H., Luo, F., Hu, H., Cheng, L., Xiao, H., Ge, R., 2020. Deeppower: Non-intrusive and deep learning-based detection of iot malware using power side channels, in: Proceedings of the 15th ACM Asia conference on computer and communications security, pp. 33–46.

He, H., Maheshwari, P., Pommerenke, D.J., 2016. The development of an em-field probing system for manual near-field scanning. IEEE transactions on Electromagnetic Compatibility 58, 356–363.

Khan, H.A., Sehatbakhsh, N., Nguyen, L.N., Callan, R.L., Yeredor, A., Prvulovic, M., Zajić, A., 2019a. Idea: Intrusion detection through electromagnetic-signal analysis for critical embedded and cyber-physical systems. IEEE Transactions on Dependable and Secure Computing 18, 1150–1163.

Khan, H.A., Sehatbakhsh, N., Nguyen, L.N., Prvulovic, M., Zajić, A., 2019b. Malware detection in embedded systems using neural network model for electromagnetic side-channel signals. Journal of Hardware and Systems Security 3, 305–318.

Korac, L., et al., 2024. Ransomware: Analysis and evaluation of live forensic techniques and the impact on linux based iot systems. arXiv preprint arXiv:2403.17571 .

Nazari, A., Sehatbakhsh, N., Alam, M., Zajic, A., Prvulovic, M., 2017. Eddie: Em-based detection of deviations in program execution, in: Proceedings of the 44th Annual International Symposium on Computer Architecture, pp. 333–346.

Patsakis, C., Arroyo, D., Casino, F., 2024. The malware as a service ecosystem, in: Malware: Handbook of Prevention and Detection. Springer, pp. 371–394.

Pham, D.P., Marion, D., Mastio, M., Heuser, A., 2021. Obfuscation revealed: Leveraging electromagnetic signals for obfuscated malware classification, in: Proceedings of the 37th Annual Computer Security Applications Conference, pp. 706–719.

Riley, R.A., Graham, J.T., Fuller, R.M., Baldwin, R.O., Fisher, A., 2019. A new way to detect cyberattacks: Extracting changes in register values from radio-frequency side channels. IEEE Signal Processing Magazine 36, 49–58.

Sehatbakhsh, N., Nazari, A., Alam, M., Werner, F., Zhu, Y., Zajic, A., Prvulovic, M., 2019. Remote: Robust external malware detection framework by using electromagnetic signals. IEEE Transactions on Computers 69, 312–326.

Singh, B., Evtyushkin, D., Elwell, J., Riley, R., Cervesato, I., 2017. On the detection of kernel-level rootkits using hardware performance counters, in: Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security, pp. 483–493.

SunFounder, 2022. LCM1602C LCD Display User Manual. Available: https://sunfounder.com.

Sutherland, S., 2013. 10 evil user tricks for bypassing anti-virus. Accessed: 2026-02-19.

Wang, X., Zhou, Q., Harer, J., Brown, G., Qiu, S., Dou, Z., Wang, J., Hinton, A., Gonzalez, C.A., Chin, P., 2018. Deep learning-based classification and anomaly detection of side-channel signals, in: Cyber Sensing 2018, SPIE. pp. 37–44.

Zhu, Y., et al., 2022. Devils in the clouds: An evolutionary study of telnet bot loaders. arXiv preprint arXiv:2211.14790 .

Statements and Declarations

Ethics declaration

Ethics declaration: not applicable.

Competing interests

The authors declare that they have no competing interests.