Chapter 11
Revolutionizing Malware Detection:
A Paradigm Shift Through Optimized Convolutional Neural Networks

Marwan Omar
Capitol Technology University, USA & Illinois Institute of Technology, USA

ABSTRACT

Cybercrime has grown into a multi-billion dollar industry in recent years. Malware deployment is one of these cybercrimes' most common aspects. This malicious software has shown its ability to cripple large commercial organizations and collect significant financial tolls up to billions of dollars yearly. It targets a variety of industries, sectors, corporations, and even individual entities without discrimination. Malware writers continuously hone their techniques and raise the bar on their sophistication, creating difficult-to-detect malware that may be left unnoticed in the background for extended periods of time to get around security measures. The first accuracy rate of the baseline model is a phenomenal 98%. The accuracy of the CNN model increases to an astonishing 99.183% by increasing its complexity, outperforming the performance of the bulk of CNN models reported in the literature. This CNN model is used to forecast the appearance of new malware samples in our dataset, further demonstrating its effectiveness.

INTRODUCTION

The use of information technology has benefited modern life by enabling previously unheard-of breakthroughs in lifestyle and professional endeavors. But these developments have also brought about serious challenges and vulnerabilities. Even seemingly harmless behaviors, like visiting a malicious website or downloading an email attachment, can cause havoc and impair the operations of contemporary businesses. Neglecting to perform periodic system upgrades or unintentionally installing malicious software might leave computer systems completely vulnerable to the dangers and risks of cyberattacks. Evidently, cybercrime has increased recently, with hackers expertly undermining important companies or sectors by manipulating entire commercial entities through the use of malware (Yi, et al, 2017).

DOI: 10.4018/979-8-3693-1906-2.ch011

Notably, ransomware, a type of malware, has become a popular technique for cyber attackers, allowing them to seize control of their targets' computer systems until a ransom is paid. One of the first known ransomware incidents took place in 1989 when attendees of the International AIDS conference received malware-filled floppy disks that later restricted their access to files, according to historical precedent. A specific amount had to be delivered to a predetermined PO Box in Panama in order to recover access (US Senate, 2022). The ransomware assault landscape has changed dramatically recently as cybercriminals strategically target high-value targets, such as particular businesses that own highly sensitive data or huge financial resources. As demonstrated by the 2021 ransomware attack on Colonial Pipeline, the largest refined products pipeline in the United States, the operation of these entities is crucial to a country's economy. These attacks spread the myth that cybercrime has grown to be a multibillion dollar industry. A large part of cybercrimes involve the use of different malware types. Nevertheless, as antivirus technologies develop into anti-malware software, malware developers also innovate and design increasingly complex and effective iterations, typified by stealthiness and covert persistence that defeat traditional security measures (Kaspersky, 2019).

In recent years, there have been steadily more malware samples discovered in the wild. Notably, research by McAfee laboratories revealed that 7,899 new, distinctive hashes were among the 1,224,628 malware threats that were discovered in the fourth quarter of 2020 (McAfee, 2020). Given that cyber attackers are always creating new malware variants, this evolving threat landscape emphasizes the necessity of strengthening malware detection and protection (Cybersecurity Ventures, 2018).

Before appropriate steps, such as isolation and quarantine, can be implemented, the classification of malware is a crucial step in recognizing and understanding the nature of malware. The two main methodologies used in this categorization process are the behavior-based approach and the signature-based approach. The former has traditionally been successful because of its accuracy and speed, but it struggles to identify malware versions that use obfuscation techniques including packing, encryption, metamorphism, and polymorphism (Souri, 2018). The behavior-based classification method overcomes the drawbacks of signature-based classification because the behavioral characteristics of different malware strains are very similar. However, gathering information on malware behaviors takes a lot of time because it must be done during malware activity.

Recent years have seen the emergence of an innovative method for classifying malware that makes use of image processing techniques (Li, 2020). By analyzing the malware's image textures, the classifier is able to recognize and categorize the malware using this method. This methodology overcomes the limitations of existing methods that rely on signature- or behavior-based analysis, strengthening malware classification techniques. Malware analysis is a step in the traditional malware detection and classification process that involves observing the actions and motives of a malicious URL or file. This analysis includes static analysis, dynamic analysis, or a combined static and dynamic analysis. While dynamic analysis observes an application's behavior while it is being used, static analysis entails extracting characteristics by looking at an application's manifest and disassembled code (Xu, 2016). Due to its ability to combat the sophisticated threats offered by the most recent malware types, hybrid analysis—which combines parts of both—emerges as the most effective strategy. The enhancement of hybrid analysis by integration with deep learning technologies was highlighted in a study by Xu et al. (2016), which produced a remarkable accuracy range of 95% to 99% in malware detection (Xu, 2016).

There has been a long-standing effort to use machine learning, particularly deep learning models, in malware detection, and cybersecurity experts are becoming more and more interested in malware visualization. For the purpose of identifying and categorizing known malware, several traditional machine