
**UNIVERSIDADE FEDERAL DE PERNAMBUCO**
**CENTRO DE TECNOLOGIA E GEOCIÊNCIAS**
**DEPARTAMENTO DE ELETRÔNICA E SISTEMAS**
**GRADUAÇÃO EM ENGENHARIA ELETRÔNICA**

**LEONARDO SILVINO BRITO**

# ANTIVIRUS DE ÚLTIMA GERAÇÃO BASEADO NA DETECÇÃO DE EXTENSÕES DO GOOGLE CHROME MALICIOSAS COM O USO DE REDES NEURAIS EXTREMAS

**RECIFE**

**2022**

**UNIVERSIDADE FEDERAL DE PERNAMBUCO**
**CENTRO DE TECNOLOGIA E GEOCIÊNCIAS**
**DEPARTAMENTO DE ELETRÔNICA E SISTEMAS**
**GRADUAÇÃO EM ENGENHARIA ELETRÔNICA**

**LEONARDO SILVINO BRITO**

# ANTIVIRUS DE ÚLTIMA GERAÇÃO BASEADO NA DETECÇÃO DE EXTENSÕES DO GOOGLE CHROME MALICIOSAS COM O USO DE REDES NEURAIS EXTREMAS

TCC apresentado ao Curso de Engenharia Eletrônica da Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, como requisito para a obtenção do título de engenheiro eletrônico.

**Orientador:** Prof. Dr. Sidney Marlon Lopes de Lima.

**RECIFE**

**2022**

Ficha de identificação da obra elaborada pelo autor,
através do programa de geração automática do SIB/UFPE

Silvino Brito, Leonardo .
Antivirus de última geração baseado na detecção de extensões do Google Chrome Maliciosas com o uso de redes neurais extremas / Leonardo Silvino Brito. - Recife, 2023.
60 : il., tab.

Orientador(a): Sidney Marlon de Lopes Lima
Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, Engenharia Eletrônica - Bacharelado, 2023.

1. Redes Neurais Artificiais. 2. Segurança de Sistemas. 3. Mapeamento de Caracteristicas. 4. Computação Forense. I. Lopes Lima, Sidney Marlon de. (Orientação). II. Título.

000 CDD (22.ed.)

**Universidade Federal de Pernambuco**
**Departamento de Engenharia Eletrônica**
**Centro de Tecnologia e Geociências**

## ATA DE SESSÃO DE DEFESA DE TRABALHO DE CONCLUSÃO DE CURSO

Aos 8 dias do mês de Novembro do ano de Dois Mil e Vinte e Dois, às 09:00, por meio de videoconferência (ONLINE), reuniu-se a banca examinadora para a sessão pública de defesa do Trabalho de Conclusão de Curso em Engenharia Eletrônica da Universidade Federal de Pernambuco, intitulado: "Antivirus de última geração baseado na detecção de extensões do Google Chrome Maliciosas com o uso de redes neurais extremas", elaborado pelo (a) aluno (a) Leonardo Silvino Brito, matrícula 11386981486, composta pelos professores Sidney Marlon Lopes de Lima e Raul Camelo DE ANDRADE ALMEIDA JUNIOR (membro titular). Após a exposição oral, o (a) candidato (a) foi arguido (a) pelos componentes da banca que em seguida reuniram-se reservadamente e deliberaram pela aprovação do candidato, atribuindo-lhe a média dez (10,0), julgando-o (a) apto (a) à conclusão do curso de Engenharia Eletrônica. Para constar, redigi a presente ata aprovada por todos os presentes, que vai assinada por mim e pelos demais membros da banca.

Prof.(a)/Membro: Prof. **Sidney Marlon Lopes de Lima** Nota: 10,0

Assinatura Documento assinado digitalmente gov.br SIDNEY MARLON LOPES DE LIMA Data: 08/11/2022 15:07:30-0300 Verifique em https://verificador.iti.br

Prof.(a)/Membro: Prof. **Raul Camelo DE ANDRADE ALMEIDA JUNIOR** Nota: 10,0

Assinatura

Recife, 8 de Novembro de 2022.

Documento assinado digitalmente gov.br GUILHERME NUNES MELO Data: 21/11/2022 10:53:53-0300 Verifique em https://verificador.iti.br

Prof(a). Guilherme Nunes de Melo
Coordenador(a) do Curso de Engenharia Eletrônica

# Resumo

Apesar da presença massiva dos antivírus nos computadores pessoais, as extensões maliciosas dos navegadores continuam atuando. Uma das razões desse insucesso é porque os antivírus não consideram as extensões como aplicativos. Portanto costuma não haver a auditoria das extensões suspeitas pelos antivírus. Atualmente, ao invés de infecções convencionais através de arquivos executáveis portáveis, os *malware* modernos empregam as extensões dos navegadores. O presente trabalho tem como objetivo a criação de um software antivírus visando detectar extensões maliciosas do Google Chrome (CRX). A amostra suspeita CRX é executada para infectar, propositalmente, o sistema operacional Windows monitorado em um ambiente controlado. Ao todo, nosso antivírus monitora e considera 1.098 ações que o arquivo CRX suspeito pode realizar quando executado. Os comportamentos auditados servem com neurônios de entrada das redes neurais autorais. O objetivo é reconhecer o padrão de *add-ons* maliciosos e separá-los dos benignos. Ao invés de redes profundas, as redes autorais são de baixa complexidade computacional. Devido aos excelentes resultados em diferentes áreas, criou-se o senso comum de que o aprendizado profundo sempre pode fornecer os melhores resultados. Na verdade, essa consideração é falsa. Para provar o embasamento teórico o antivírus autoral emprega redes neurais morfológicas superficiais ao invés de redes convolucionais profundas. O antivírus autoral tem sua acurácia comparada com os antivírus de última geração baseados em redes neurais superficiais e profundas. O antivírus autoral pode combinar alta acurácia com tempo de aprendizado reduzido. O antivírus criado alcançou um desempenho médio de 99,99% na distinção entre arquivos CRX benignos e *malware*. O treinamento demanda um tempo médio de 0,60 segundos. Diferentes condições iniciais, funções de aprendizado e arquiteturas do antivírus são investigadas. As limitações dos antivírus comerciais podem ser supridas por antivírus inteligentes. Ao invés de modelos baseados em lista negra, o antivírus autoral detecta *malware* CRX de forma preventiva e não reativa, ao contrário do *modus operandi* dos antivírus tradicional.

**Palavras-chave**:

Redes Neurais Artificiais; Segurança de Sistemas; Mapeamento de Caracteristicas; Computação Forense.

# Abstract

Despite the massive presence of antivirus software on personal computers, malicious browser extensions are still at work. One of the reasons for this failure is because antiviruses do not consider extensions to be applications. Thus there is usually no auditing of suspicious extensions by antiviruses. Nowadays, instead of conventional infections via portable executable files, modern malware employs browser extensions. The present work aims to create antivirus software. It uses machine learning and artificial intelligence to detect malicious Google Chrome extensions. The suspected CRX (Google Chrome's extension) sample is executed to infect the Windows OS monitored in a controlled environment. In all, our antivirus monitors and considers, 1,098 actions that the suspected CRX file can perform when executed. The audited behaviors serve with input neurons of the authorial neural networks. The goal is to recognize the pattern of malicious add-ons and segregate them from benign ones. Instead of deep networks, our networks are of low computational complexity. Due to excellent results in different areas, a common sense has been created that deep learning can always provide the best results. In fact, this consideration is false. To prove our theoretical basis, the authorial antivirus employs shallow morphological neural networks instead of deep convolutional networks. Authorial antivirus has accuracy compared to next-generation antiviruses based on shallow and deep neural networks. Our antivirus can combine high accuracy with reduced training time. Our antivirus achieved an average performance of 99.99% in the distinction between benign and malware CRX files. The training consume an average time of 0.60 seconds. Different initial conditions, learning functions and architectures of our antivirus are investigated. The limitations of commercial antiviruses can be supplied by intelligent antiviruses. Instead of blacklist-based models, our antivirus detect CRX malware preventively and not reactively. Contrary, there are traditional antivirus *modus operandi*.

**Keywords**:Artificial neural networks; Systems Security; Mapping of Characteristics; Forensic Computing.

## LISTA DE ILUSTRAÇÕES

## LISTA DE TABELAS

## LISTA DE ABREVIATURAS E SIGLAS

**CRX** Chrome Extension

**MLP** Multilayer Perceptron

**ELM** Extreme Learning Machine

**PE** Portable Executable

**LSTM** Long Short-Term Meomry

**RNN** Recurrent Neural Network

**PCAP** Packet Capture

**API** Application Programming Interface

**OS** Operational System

**mELM** morphological ELM

**fmELM** fuzzy-morphological ELM

**SUMÁRIO**

# 1 INTRODUÇÃO

As extensões do navegador Google Chrome estão presentes no dia a dia dos usuários de internet. De forma direta ou direta, as extensões permitem aumentar a produtividade, reduzir distrações e obter conteúdo, seja educacional ou profissional. As extensões são capazes de integrar o navegador a um serviço-online, em acréscimo, as extensões permitem sincronizar afazeres particulares aos serviços dos grandes provedores de dados em nuvens.

Como adversidade, algumas extensões podem conter atividades maliciosas. A aquisição de extensões mal-intencionadas geralmente ocorre através de sites avulsos, mas em alguns casos, extensões maliciosas também são disponibilizadas nas lojas oficiais como o Google Play. Por exemplo, pesquisadores da ICEBRG, firma de segurança de informação dos Estados Unidos, descobriram recentemente quatro extensões na Google Chrome *Web Store* que demonstravam serem aplicativos de lembrete inofensivos, mas que na realidade geravam receitas para seus criadores acessando furtivamente em anúncios pagos por clique (CIMPANU, 2018).

Durante um estudo de três meses, pesquisadores da *Awake Security* descobriram 111 extensões maliciosas disponíveis para download na loja oficial do Google Chrome. Setenta e nove dessas extensões estavam presentes na Google Chrome *Web Store* (VIRGILLITO, 2021). Embora a maioria das extensões aparentavam funcionar normalmente, na verdade eles estavam oferecendo suporte a uma campanha massiva de espionagem global em vários ramos industriais. Os pesquisadores também identificaram que as referidas extensões foram baixadas quase que 33 milhões de vezes por usuários do Google Chrome. Algumas dessas extensões receberam mais de dez milhões de instalações. As referidas extensões capturavam os pressionamentos de tecla (como senhas), capturas de tela e coleta de *tokens* de credenciais presentes em parâmetros ou cookies. Mesmo as extensões mais básicas geralmente exigem permissão para "ler e modificar todos os seus dados nos sites que você visita". Através da referida permissão, a extensão suspeita pode fazer praticamente o que quiser com os dados íntimos dos usuários.

De acordo com *Ponemon Institute 2015 Cost of Data Breach Study*, o custo médio de uma violação de dados é de 3,79 milhões de doláres em todo o mundo 6,5 milhões de doláres especificamente para empresas norte-americanas. No entanto as empresas em indústrias fortemente regulamentadas, como serviços de saúde, farmacêuticos e financeiros têm prejuízos substancialmente mais altos (PONEMON, 2015).

Apesar da presença massiva dos antivírus em praticamente todos os computadores pessoais, as extensões maliciosas continuam atuando. Uma das razões desse insucesso é porque os antivírus não consideram as extensões como aplicativos. Portanto, costuma não haver a auditoria das extensões suspeitas pelos antivírus. Atualmente, ao invés de infecções convencionais através de arquivos executáveis portáveis, os *malware* (malicioso + software) modernos empregam as

extensões dos navegadores.

Como adversidade, extensões maliciosas são bastante difíceis de catálogo (*VirushShare*) (VIRUSSHARE, 2022). Graças à natureza do negócio de hospedagem de sites, quando uma empresa descobre que um de seus servidores contém atividades maliciosas, é bem mais vantajoso simplesmente apagar esse servidor e configurar um novo ao invés de tentar descobrir o que aconteceu (SOPHOS, 2014). Como nem as referidas empresas nem seus parceiros de segurança sabem exatamente o que aconteceu, geralmente o novo servidor também conterá atividades maliciosas (SOPHOS, 2014). É importante atentar-se que 85% de todos os softwares maliciosos se espalham por meio de navegadores da web. Ainda mais alarmante, 94% dos malware totalmente indetectáveis são entregues via navegação na web (MCAFEE, 2017).

Em síntese, extensões maliciosas têm capacidade de ludibriar os mecanismos de *cyber*-vigilância. Então, o trabalho proposto investiga os 68 principais antivírus comerciais mundiais. A detecção de extensões maliciosas variou entre 0% a 63,64% a depender do antivírus. Em média, houve a detecção de 46,08% dos *malware*. Com aspecto desfavorável, os antivírus, em média, atestaram falsos negativos e foram omissos em 44,85% e 8,92% dos casos, respectivamente. Além disso, cerca de 7,35% dos antivírus não foram capazes de diagnosticar qualquer uma das extensões maliciosas. Enfatiza-se que no estudo, os *malware* analisados têm as suas malfeitorias documentadas e catalogadas (VIRUSSHARE, 2022). Mesmo assim, uma parcela considerável dos antivírus comerciais avaliados não tinham qualquer conhecimento sobre as existências das extensões maliciosas do Google Chrome.

Atualmente, as organizações buscam suprir as deficiências dos antivírus tradicionais através de técnicas de inteligência artificial baseadas em aprendizado de máquina. Inteligência artificial consegue automatizar muitas tarefas, analisando milhares de dados, extraindo características deles e os classificando. Logo, a inteligência artificial é capaz de reconhecer padrões de comportamento previamente classificados como suspeitos em tempo real (LIMA; SILVA; LUZ, 2021).

Redes neurais são modelos de inteligência artificial frequentemente utilizados para resolver problemas de reconhecimento de padrões tendo como principal característica o poder de generalização diante de dados não apresentados à rede. Em grande parte das redes neurais, como a Multilayer Perceptron (MLP) (Perceptron com Múltiplas Camadas), é necessário um conhecimento sobre os parâmetros da rede para obter máximo desempenho na solução do problema. Uma preocupação comum nesse tipo de rede é evitar se ater a mínimos locais, sendo necessário adicionar métodos de controle que permitam à rede desprender-se dessas regiões (de mínimos locais). Outra característica comum nesse tipo de rede é a grande quantidade de tempo de treinamento necessária para torná-la apta a realizar classificações corretamente. Apesar de excelentes acurácias, as redes neurais de última geração, especificamente *Deep Learning* (Aprendizado Profundo), podem requerer uma duração excessiva até a conclusão do seu treinamento.

Tecnicamente, em termos de inteligência artificial, quando uma nova vulnerabilidade for detectada, deve haver uma nova etapa de aprendizado (treinamento) das redes neurais artificiais empregadas pelos mecanismos de segurança digital. O(s) novo(s) atributos referente(s) à vulnerabilidade recém-detectada devem ser agrupados aos atributos convencionais, previamente catalogados. Dessa forma, é possível proteger os demais computadores, ainda não infectados, das extensões que explorem essa falha recém-encontrada. Então, quanto mais rápido for o tempo de treinamento do modelo de inteligência computacional, maiores são as chances de prevenção da infecção dos computadores pessoais e das organizações. Caso o tempo de treinamento do antivírus seja elevado, a exploração da vulnerabilidade pode corromper os web-navegadores através de extensões mal intencionadas.

O trabalho proposto aplica as redes neurais que utilizam Extreme Learning Machine (ELM)(Máquinas de Aprendizado Extremo) na área de segurança da informação, especificamente no reconhecimento de padrão de atividades maliciosas em extensões do navegador Google Chrome. As redes ELMs têm como principal característica a velocidade de treinamento e a previsão de dados. As ELMs são adequadas à Perícia Forense Digital visto que são lançados 8 (oito) novos *malware* por segundo (INTEL, 2018). Então, paradoxalmente um antivírus recém-lançado já pode estar obsoleto e necessitar de um novo treinamento mediante uma vulnerabilidade recém-descoberta. Em síntese, o tempo de aprendizado de um antivírus não deve ser discrepante em comparação à taxa de criação de novos malware mundialmente.

Ao invés de kernels convencionais, métodos que permitem aplicar classificadores lineares a problemas não lineares, o presente trabalho emprega kernels autorais para as ELMs. Os kernels são funções matemáticas utilizadas como método de aprendizado das redes neurais. O aprendizado baseado em kernel oferece a possibilidade da criação de um mapeamento não-linear de dados sem que haja a necessidade do aumento do número de parâmetros ajustáveis como, por exemplo, taxa de aprendizagem comumente empregada em redes neurais com retropropagação. Os kernels autorais são inspirados em operadores morfológicos de processamento de imagem de Erosão e Dilatação. O trabalho proposto estima que os kernels morfológicos sejam capazes de se adequar a qualquer fronteira de decisão.

Quanto aos cenários e experimentos, o antivírus autoral alcança um desempenho médio de 99.99% na distinção entre extensões benignas e maliciosas do Google Chrome. Então, o presente artigo demonstra que a inteligência artificial é uma boa alternativa para as fabricantes dos antivírus comerciais. As limitações dos mecanismos de cyber-segurança tradicionais podem ser supridas pelo antivírus autoral. Ao invés de modelos baseados em listas negras, a nossa engine emprega ciência de dados, aprendizagem de máquina e inteligência artificial na identificação de comportamentos maliciosos. O trabalho proposto tem, como destaque, as seguintes contribuições: Investigação dos principais 68 antivírus comerciais quanto à identificação de extensões maliciosas do navegador Google Chrome. Em média, houve a detecção de 16,82% dos *malware*. O antivírus autoral alcança um desempenho médio de 99,99% na distinção entre extensões benignas e

maliciosas do Google Chrome acompanhado de um tempo de treinamento médio de apenas 0,04 segundos. O antivírus criado possibilita a detecção preventiva das ameaças virtuais, em ambiente controlado, antes de alcançarem os web-navegadores dos clientes.

Este trabalho está organizado da seguinte forma. Na Seção 2, são apresentado as restrições dos antivírus empresariais. Na Seção 3, descreve sobre o estado da arte em relação aos antivírus de inteligência artificial. Na Seção 4, é apresentada a técnica proposta. Na Seção 5, é realizada uma avaliação entre a rede neural ELM autoral e as redes neurais ELM clássicas. Na Seção 6, são apresentados resultados e algumas percepções. Por fim, na Seção 7, é apresentam-se as conclusões gerais e discute-se das visões futuras do trabalho.

# 2 LIMITAÇÕES DOS ANTIVÍRUS COMERCIAIS

Embora questionado há mais de uma década, o modus operandi dos antivírus convencionais é baseado em assinaturas quando o arquivo suspeito é consultado em conjuntos de dados nomeados em uma lista negra. Portanto, basta que o *hash* do arquivo investigado não esteja na lista negra do antivírus para que o malware não seja detectado. As funções de *hash* são um identificador exclusivo para um determinado arquivo.

Dadas as limitações dos antivírus comerciais, não é uma tarefa difícil desenvolver e distribuir variantes de aplicativos maliciosos. Para isso, basta fazer pequenas alterações no malware original com rotinas que, efetivamente, não têm utilidade, como loops de repetição e ramificações condicionais sem instruções em seus escopos. Essas alterações sem utilidade, no entanto, transformam o *hash* do malware modificado do *hash* do malware original. Consequentemente, o malware aumentado com rotinas nulas não é reconhecido pelo antivírus que reconhece o malware inicial. Ressalta-se a existência de exploits responsáveis por criar e distribuir, de forma automatizada, variantes do mesmo malware original. Conclui-se que os antivírus, baseados em assinaturas, têm eficácia nula quando submetidos a variantes do mesmo software (SANS, 2017)(LIMA, 2020).

Através da plataforma VirusTotal, este trabalho proposto explora 68 antivírus comerciais com seus respectivos resultados apresentados na Tabela 1. Foram utilizados 22 extensões maliciosas do Google Chrome obtidas do VirusShare (BASE, 2022). O objetivo do trabalho é verificar o número de amostras maliciosas catalogadas por antivírus. A motivação é que a aquisição de novos malwares é primordial para combater atividades maliciosas.

Quanto maior o conjunto de dados, denominado lista negra, melhor tende a ser a defesa dada pelo antivírus. Primeiro, o malware Chrome Extension (CRX) (tipo de arquivo das extensões do Chrome) é enviado para o servidor pertencente à plataforma VirusTotal. Neste ponto, os arquivos CRX foram analisados pelos 68 antivírus comerciais do VirusTotal. Em seguida, os antivírus fornecem seus diagnósticos para amostras CRX enviadas ao servidor. O VirusTotal permite que três tipos diferentes de diagnóstico sejam emitidos: benigno, malware e omissão.

Para a primeira possibilidade do VirusTotal, o antivírus detecta a malícia do arquivo suspeito. Dentro do ambiente experimental proposto, todas as amostras enviadas são malwares documentados por respondentes de incidentes. A detecção de malware mostra que o antivírus oferece suporte robusto contra invasões digitais. Na segunda possibilidade, o antivírus certifica a benignidade do arquivo definido. Então, no estudo proposto, quando o antivírus afirma que o arquivo é benigno, trata-se de um falso negativo, pois todas as amostras enviadas são maliciosas. Em outras palavras, o arquivo investigado é malware; no entanto, o antivírus erroneamente atesta que é benigno. Dentro da terceira possibilidade, o antivírus não fornece uma análise do aplicativo

suspeito. A omissão mostra que o arquivo investigado nunca foi avaliado pelo antivírus, portanto, não pode ser avaliado em tempo real. A omissão do diagnóstico por antivírus aponta para sua limitação em serviços de grande porte.

A Tabela 1 mostra os resultados dos 68 produtos antivírus avaliados. O antivírus Kaspersky obteve o melhor desempenho ao detectar 63,64% dos malwarse investigados. Uma das maiores adversidades na combinação de aplicativos maliciosos é o fato de os fabricantes de antivírus não compartilharem suas listas negras de malware devido a disputas comerciais. Por meio da análise da Tabela 1, o trabalho proposto aponta para um agravante dessa vantagem: o mesmo fabricante de antivírus não compartilha seus bancos de dados entre seus diferentes antivírus. Observe, por exemplo, que os antivírus McAfee-GW-Edition e McAfee pertencem à mesma empresa. Suas listas negras, embora robustas, não são compartilhadas entre si. Portanto, as estratégias comerciais de uma mesma empresa atrapalham o enfrentamento aos malwares, o que demonstra que os fabricantes de antivirais não estão necessariamente preocupados em evitar invasões cibernéticas, mas sim em otimizar o faturamento de seus negócios.

A identificação de malware variou de 0% a 63,64%, a depender do antivírus. Em média, os 68 antivírus identificaram 34,95% dos malwares examinados, com desvio padrão de 40,92%. O desvio padrão elevado mostra que o reconhecimento de arquivos maliciosos pode mudar abruptamente dependendo do antivírus escolhido. A proteção contra intrusões digitais está na função de escolher um antivírus vigoroso com uma lista negra expansiva e atualizada. No geral, os antivírus certificaram falsos negativos em 67,82% dos casos, com desvio padrão de 33,87%. A atenção à benignidade do malware pode estar implicada em danos irrecuperáveis. Uma pessoa ou instituição, por exemplo, pode começar a confiar em um determinado aplicativo malicioso quando, na verdade, é um malware. Em média, os antivírus foram omitidos em 8,07% dos casos, com desvio padrão de 21,17%. A omissão do diagnóstico foca na restrição do antivírus em reconhecer malware em tempo real.

Devido à dificuldade de combater aplicativos maliciosos, os antivírus comerciais não possuem um padrão na classificação de malware, conforme encontrado na Tabela 2. Foram escolhidos 3 dos 22 malwares CRX para exemplificar as diversas classificações dadas pelas atividades antivirais comerciais. Como não há padrão, os antivírus usam os nomes que quiserem; por exemplo, uma empresa pode identificar o malware CRX como "Android: RuFraud-I"e uma segunda empresa pode identificá-lo como "Artemis! 9EF6966B98A5". Portanto, a falta de um padrão atrapalha as estratégias de segurança cibernética, pois cada categoria de malware deve ter tratamentos diferentes (vacinas). Conclui-se que é inviável que o aprendizado de máquina supervisionado adote o reconhecimento de padrões para as categorias de malware CRX. Devido a esse emaranhado confuso de classificação multiclasse fornecida por especialistas (antivírus), como visto na Tabela 2, é estatisticamente improvável que qualquer técnica de aprendizado de máquina adquira capacidade de generalização.

Tabela 1 – Resultados de antivírus comerciais. Resultados expandidos de 68 antivírus comerciais em todo o mundo estão no repositório autoral (BASE, 2022).

| **Antivirus** | Detecção (%) | Falso negativo (%) | Omissão (%) |
|---|---|---|---|
| Kaspersky | 63,64 | 36,36 | 0 |
| DrWeb | 59,09 | 40.91 | 0 |
| Jiangmin | 54,55 | 45,45 | 0 |
| ESET-NOD32 | 54,55 | 4,45 | 0 |
| Antiy-AVL | 54,55 | 40,91 | 4,55 |
| NANO-Antivirus | 54,55 | 45,45 | 0 |
| BitDefender | 54,55 | 45,45 | 0 |
| Emsisoft | 54,55 | 45,45 | 0 |
| ZoneAlarm | 50 | 50 | 0 |
| Avast | 50 | 50 | 0 |
| Baidu | 0 | 100 | 0 |
| ClamAV | 0 | 90,91 | 9,09 |
| SUPERAntiSpyware | 0 | 100 | 0 |
| Ad-Aware | 0 | 100 | 0 |
| Sophos | 0 | 100 | 0 |
| Kingsoft | 0 | 100 | 0 |
| Gridinsoft | 0 | 50 | 50 |
| ViRobot | 0 | 100 | 0 |
| Cynet | 0 | 50 | 50 |
| AVware | 0 | 4,55 | 95,45 |

Fonte:O autor (2022).

Tabela 2 – Resultado da submissão de três malwares ao VirusTotal. Resultados expandidos de 68 antivírus comerciais em todo o mundo estão no repositório autoral (BASE, 2022).

| **Antivirus** | $VirusShare_A$ | $VirusShare_B$ | $VirusShare_C$ |
|---|---|---|---|
| ALYac | Trojan.Linux.Generic | Gen:Variant.Backdoor.Linux | Trojan.Linux.GenericA |
| AVG | ELF:DDoS-S [Trj] | ELF:Gafgyt-DO [Trj] | ELF:DDoS-S [Trj] |
| Ad-Aware | Trojan.Linux.Generic | \|Gen:Variant.Backdoor.Linux | Trojan.Linux.GenericA |
| AhnLab-V3 | Linux/Mirai.Gen6 | Linux/Tsunami.Gen | Linux/Gafgyt.Gen |
| Antiy-AVL | Trojan[Backdoor]/Linux | GrayWare/Linux.Generic | Trojan[Backdoor]/Linux |
| Arcabit | Trojan.Linux.Genericr | Trojan.Backdoor | Undected |
| Avast | ELF:DDoS-S [Trj] | ELF:Gafgyt-DO [Trj] | ELF:DDoS-S [Trj] |
| Avast-Mobile | ELF:DDoS-S [Trj] | ELF:Tsunami-CR [Trj] | ELF:DDoS-S [Trj]\| |
| Avira | LINUX/Gafgyt.opnd | LINUX/Tsunami | Undected |
| Baidu | Undetected | Undetected | Undetected |
| BitDefender | Trojan.Linux.Generic | Gen:Variant.Backdoor.Linux | Trojan.Linux.GenericA |
| BitDefenderTheta | Gen:NN.Mirai.34608 | Gen:NN.Mirai.34608 | Gen:NN.Mirai.34608 |
| Bkav | Undetected | Undetected | Undetected |
| CAT-QuickHeal | Undetected | Elf.Trojan.A1198970 | Undetected |
| CMC | Undetected | Undetected | Undetected |
| ClamAV | Unix.Malware.Agent | Unix.Trojan.Mirai | Unix.Trojan.Gafgyt-111 |
| MAX | malware (ai score=99) | Undetected | malware (ai score=98) |
| MaxSecure | Trojan.Malware | Linux/Gafgyt.h | Trojan.Malware |
| McAfee | Linux/Gafgyt.h | Linux/Gafgyt.r | Undected |
| Kaspersky | HEUR:Backdoor.Linux.Gafgyt | HEUR:Backdoor.Linux.Tsunami | HEUR:Backdoor.Linux.Gafgyt. |
| Rising | Backdoor.Mirai | Undetected | Backdoor.Gafgyt/Linux\| |

Fonte:O autor (2022).

# 3 ESTADO DA ARTE

Um arquivo CRX é uma extensão que adiciona funcionalidade adicional ou design ao web-navegador Google Chrome. O arquivo é salvo em formato compactado e pode conter arquivos JavaScript, .JSON além de outros arquivos como imagens e aplicativos executáveis. Os arquivos CRX são usados para instalar *plug-ins* de navegador, tais como jogos, bloqueadores de anúncios e leitores de notícias. A importância das extensões não é apenas para entretenimento, mas também para educação, e até mesmo para o trabalho. Visando proteger o usuário das extensões maliciosas do Google Chrome, o antivírus autoral tem como alvo malware CRX.

O *modus operandi* do software antivírus comercial é usado para identificar malware CRX com base em assinatura. O principal problema dessa estratégia é que, para que um novo malware seja assinado(lista negra dos antivírus comerciais), é preciso detectar que determinados computadores foram infectados. Considerando as limitações dos softwares antivírus comerciais e gerenciadores de segurança, a mais recente tecnologia propõe extrair e analisar arquivos por meio de uma máquina de aprendizado estatístico. A inteligência artificial pode automatizar muitas tarefas analisando milhares de arquivos, extraindo suas funções e classificando-os.

A Tabela 3 apresenta uma lista de alguns antivírus importantes do estado da arte. Como os antivírus de última geração visam sistemas distintos (Android, Windows). Todos esses algoritmos de aprendizado foram replicados usando os métodos mostrados neste trabalho para evitar comparações injustas. A comparação entre o antivírus autoral e o de última geração está na Seção 7.

LIMA, *et al.*(2021) criou um antivírus capaz de detectar malware em arquivos da extensão Portable Executable (PE) com uma acurácia média de 98,32% (LIMA; SILVA; LUZ, 2021). LIMA, *et al.*(2021) usou quatro arquiteturas para cada função de aprendizado MLP. A primeira arquitetura tem uma única camada oculta contendo 100 neurônios. A segunda arquitetura possui duas camadas ocultas, cada uma com 100 neurônios. A terceira arquitetura emprega uma camada oculta; no entanto, 500 neurônios são usados. A quarta arquitetura possui quatro camadas ocultas, cada uma com 1.261 neurônios. Dentro da quarta arquitetura, o número de 1.261 neurônios é porque a fórmula empírica de Hecht-Nielsen é aplicada. Hecht-Nielsen (1987) observou que a rede neural pode ser configurada como uma camada oculta com exatamente 2n+1 nós, onde n é o número de nós de entrada. Os antivírus feitos por LIMA, *et al.*(2021) usavam redes neurais rasas.

Os antivírus baseados em rede profunda também alcançaram excelente acurácia. VINAYAKUMAR, R. *et al.* (2019) alcançou uma acurácia média de 98,90% na detecção de malwares de arquivos da extensão PE (VINAYAKUMAR; SOMAN, 2018). A estrutura de rede profunda tem 34 camadas [1]. A rede é treinada com 500 épocas com um tamanho de lote de

[1] https://github.com/vinayakumarr/dnn-ember/blob/master/DNN-info.pdf

treinamento de 64 e uma taxa de aprendizado de 0,01

Tabela 3 – Resumo das principais técnicas de antivírus de última geração.

| Autores | Tipo de Rede Neural | Técnina de Rede Neural | Dispositvo | Acurácia |
|---|---|---|---|---|
| Antivírus Autoral | Rede rasa | Morphological ELM (mELM) | *Desktop* | 99,99% |
| WOZNIAK, M, *et al,* (2015) | Rede profunda | Long-Short Term Memory (LSTM) | *Firewall* | 99,99% |
| FARUKI, P, *et al,* (2019) | Rede profunda | Rectified Linear Unit (ReLU) | *Android* | 98,65% |
| VINAYAKUMAR, R,*et al,* (2019) | Rede profunda | DeepMalNet | *Desktop* | 98,90% |
| LIMA, *et al,* (2021) | Rede rasa | Multi-Layer Perceptron (MLP) | *Desktop* | 98,32% |
| HARDY, W, *et al,* (2016) | Rede profunda | Stacked Autoencoders | *Desktop* | 96,85% |
| MANIATH, S, *et al,* (2017) | Rede profunda | Long-Short Term Memory (LSTM) | *Desktop* | 96,67% |
| HOU, S, *et al,* (2016) | Rede profunda | Deep Belief | *Android* | 96,66% |
| SU, *et al,* (2018) | Rede profunda | Training batch (CNN) | *IoT* x86 | 94,00% |

Fonte:O autor (2022).

SU, J. *et al.* (2018) alcançou uma acurácia média de 94,00% na detecção de malware da Internet das Coisas (IoT) (SU; VASCONCELLOS D., 2018). A estrutura de rede profunda tem 6 camadas. Existem 3 camadas com pesos apreensíveis: 2 camadas convolucionais e 1 camada totalmente conectada. A rede é treinada com 5.000 iterações com um tamanho de lote de treinamento de 32 e uma taxa de aprendizado de 0,0001.

O modelo de rede neural recorrente foi aplicado à detecção de malware de computadores pessoais para a Internet das Coisas (MANIATH; ASHOK, 2017)(WOZNIAK; SILKA, 2015). MANIATH, S. *et al.* (2017) criou um software antivírus para detectar ransomware empregando redes profundas Long Short-Term Meomry (LSTM) (MANIATH; ASHOK, 2017). LSTM é uma arquitetura de

# 4 MATERIAIS E MÉTODOS

O presente trabalho tem como objetivo elaborar a recuperação de arquivos CRX (Extensões do Google Chrome) submetidos à análise dinâmica. Trata-se de um conjunto de dados que permite a classificação de arquivos com a extensão CRX entre benignos e malware. O conjunto de dados autoral consiste em 22 amostras de arquivos CRX de malware e 990 outras amostras CRX benignas. O conjunto de dados autoral seria inadequado para aprendizado dotado de inteligência artificial (IA), considerando que os arquivos CRX apresentaram quantidades desbalanceadas nas diferentes classes (malware e benigno). O objetivo é que classificadores que são tendenciosos para uma determinada classe não tenham suas taxas de sucesso favorecidas.

Se não houvesse tratamento em nosso conjunto de dados, haveria uma tendência para maiores acertos na classe majoritária (benigno) e alta taxa de erro na classe minoritária (malware). A explicação é que o número de amostras de benignos e malware são desiguais; 22 e 990, respectivamente. Assim, quando se emprega um conjunto de dados desbalanceado, as taxas de acertos dos classificadores podem ser favorecidas. Bastaria que o classificador fosse tendencioso para a classe majoritária.

O trabalho proposto emprega uma estratégia inspirada em trabalhos de engenharia biomédica. Na área da saúde, a presença de uma anormalidade (por exemplo, câncer) ocorre a cada milhares de diagnósticos de pacientes saudáveis. A estratégia biomédica consiste em repetir o treinamento de acordo com a relação entre a classe majoritária e minoritária (22:990 = 45 iterações). No presente trabalho, a cada iteração, um novo pacote da classe majoritária é apresentado à classe minoritária (22:22). Isto garante que os classificadores tendenciosos não sejam favorecidos. Isto também mantém a diversidade de amostras da classe majoritária, contidas em nosso conjunto de dados.

Ainda quanto às precauções metodológicas tomadas pela engenharia biomédica de última geração. Reserva-se quantidades relevantes de amostras benignas e malware nos pacotes separadas para treinamento e testes. Hipoteticamente, assumindo um pacote reservado para testes com pouca ou nenhuma instância da classe de malware. Nessa situação, um classificador tendencioso para a classe benigna teria sua taxa de acerto favorecida. Há o cuidado metodológico de selecionar exemplares aleatórios, em mesma quantidade de benignos e maliciosos para os pacotes destinados a treinamento e testes.

Em relação aos vírus virtuais, foram extraídos arquivos CRX maliciosos do VirusShare, que é um repositório de amostras de malware para fornecer aos pesquisadores de segurança,respondentes, analistas forenses e o acesso mórbido curioso a amostras de código malicioso em tempo real (VIRUSSHARE, 2022). Para catalogar as 22 amostras de malwares de extensão CRX, foi necessário adquirir e analisar, por script autoral, aproximadamente 3 milhões

de amostras de malware dos relatórios atualizados diariamente pelo VirusShare.

Com relação aos arquivos CRX benignos, o catálogo foi obtido através da Chrome Web Store. Todos os arquivos benignos foram auditados pelo VirusTotal. Então, os arquivos CRX benignos contidos no autoral tiveram sua benevolência atestada pelos principais antivírus comerciais do mundo. Os resultados obtidos correspondentes às análises dos arquivos CRX benignos e malware resultantes da auditoria VirusTotal estão disponíveis para consulta no endereço virtual autoral (BASE, 2022). O objetivo de criar o conjunto de dados autoral é dar total possibilidade para que a metodologia proposta seja replicada por outros em trabalhos futuros. Então, autoral está disponível gratuitamente com todas as suas amostras, como malware benigno:

- Auditorias do VirusTotal;
- Análise dinâmica do Cuckoo Sandbox.

A autoral também disponibiliza em seu endereço virtual e seus 990 arquivos CRX benignos. Além disso, o conjunto de dados exibe a lista de todos os outros 22 arquivos CRX, desta vez malware. Em seguida, existe a possibilidade de adquirir todo o malware empregado pela autoral estabelecendo acordos e submetendo-se aos termos de uso do ViruShare (BASE, 2022). Concluí-se que o conjunto de dados autoral proporciona transparência e imparcialidade à pesquisa e demonstra a veracidade dos resultados exibidos e discutidos. Portanto, espera-se que o autoral sirva de base para a criação de novos trabalhos científicos visando antivírus de próxima geração

Todos os experimentos foram realizados em um computador equipado com 250 GB de RAM, 8 processadores e 300GB de armazenamento em massa. Portanto, não há comparações injustas, e antivírus de última geração são treinados e testados no mesmo computador usado pelo antivírus protegido por direitos autorais. O antivírus autoral requer baixa capacidade de processamento e armazenamento. Ressalta-se que o antivírus autoral pode ser utilizado em qualquer computador desktop convencional.

# 5 METODOLOGIA PROPOSTA

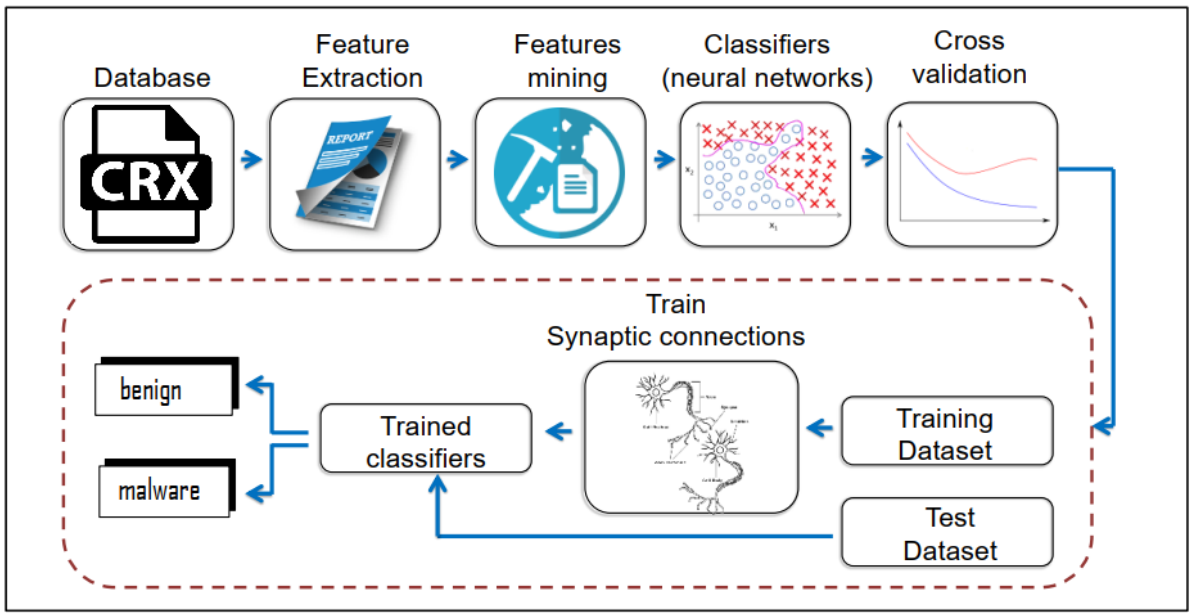
A Figura 1 mostra o diagrama de bloco da metodologia proposta. Inicialmente, o arquivo CRX, originário da base de dados é executado para verificar a tentativa de corromper a máquina virtual e, na sequência, o Windows 7 auditado pelo Cuckoo Sandbox. Os recursos dinâmicos são sintetizados na Seção 5.1. Em seguida, as características dinâmicas dos arquivos CRX são armazenadas em um formato de repositório de aprendizado de máquina.

Como método de mineração de recursos, alguns comportamentos auditados pelo sandbox são ignorados. O critério de mineração adotado refere-se à eliminação de recursos que dizem respeito a um único arquivo CRX, por exemplo, IDs de processos, nomes de processos, md5 e sha. Após a mineração de recursos, os comportamentos relevantes das amostras CRX servem como atributos de entrada de aprendizado de máquina, especificamente, as redes neurais artificiais são empregadas como classificadores. O objetivo é agrupar as amostras CRX em duas classes: benigno e malware. A etapa de classificação é explicada em detalhes na Seção 5.2. Os resultados da classificação são descritos no Capítulo 6.

## 5.1 Extração de recursos dinâmicos

Os recursos dos arquivos CRX são derivados da análise dinâmica de arquivos suspeitos. Assim, em nossa abordagem, o malware é executado para infectar intencionalmente o Windows 7 em tempo real a partir do Cuckoo Sandbox (CUCKOO, 2020). Um total de 1.098 atributos

Figura 1 – Diagrama da metodologia proposta.




Fonte:O autor (2022).

relevantes para monitorar arquivos suspeitos no ambiente controlado proposto são gerados a partir de cada arquivo CRX, incluindo os arquivos benignos, pois eles também podem realizar ações suspeitas, assim a rede neural pode melhor identificar, após treinada, se um arquivo é ou não malicioso. Para facilitar o entendimento dos neurônios da camada de entrada, o repositório autoral estende a descrição das propriedades verificadas pelo antivírus do autor. (BASE, 2022).

Em seguida, os grupos de recursos são detalhados:

- Recursos relacionados a máquinas virtuais. O objetivo é verificar se o arquivo auditado busca detectar se as máquinas virtuais Bochs, Sandboxie, VirtualBox, VirtualPC, VMware, Xen ou Parallels estão sendo utilizadas por meio da presença de chaves de registro (regedit), arquivos, instruções e drivers de dispositivos utilizados por eles.
- Recursos relacionados a malware. Verifica se o arquivo auditado tenta criar Mutexes (arquivos de nome único, com uma função para definir um estado de bloqueio/desbloqueio, o que garante que apenas um processo por vez use os recursos).
- Recursos relacionados ao Bitcoin. Examina se o arquivo testado tenta instalar a biblioteca OpenCL, ferramenta de mineração Bitcoins.
- Funcionalidades relacionadas a bots (máquinas que realizam tarefas de rede automáticas, maliciosas ou não, sem o conhecimento de seus proprietários).
- Recursos relacionados a navegadores. Verifica se o arquivo tentou:
  - modificar as configurações de segurança do navegador;
  - modificar a página inicial do navegador;
  - adquirir informações privadas de navegadores de internet instalados localmente.
- Recursos relacionados a firewalls. A análise forense digital proposta audita se o arquivo tenta modificar as políticas e configurações do firewall local.
- Recursos relacionados à computação em nuvem. O arquivo é auditado quando tenta se conectar aos serviços de armazenamento ou arquivos do Dropbox, Google, MediaFire, MegaUpload, RapidShare, Cloudflare e WeTransfer.
- Recursos que buscam desabilitar recursos do sistema operacional Windows 7 e outros utilitários. A auditoria verifica para determinar se o arquivo tenta:
  - modificar as políticas do sistema para impedir o lançamento de aplicativos ou executáveis específicos;
  - desativar avisos de segurança do navegador;
  - desabilitar recursos e propriedades de segurança do Windows;

- Recursos associados ao tráfego de rede no sistema operacional Windows 7 na extensão Packet Capture (PCAP).
- Recursos relacionados ao sistema operacional Windows 7 (Regedit):
  - Alterações nas associações entre extensões de arquivo e software instalado na máquina;
  - Alterações nas informações do usuário atual;
  - Corrupção do driver;
  - Alterações nas configurações de aparência do Windows e configurações feitas pelos usuários, como papel de parede, protetor de tela e temas;
  - Alterações nas configurações de hardware.
- Recursos relacionados ao uso de sandboxes. A perícia digital examina se o arquivo tentou desativar as funções do Windows monitoradas pelo Cuckoo Sandbox.
- Funcionalidades relacionadas a ransomware (um tipo de malware que, por meio de criptografia, deixa os arquivos da vítima inutilizáveis e, em seguida, solicita um resgate em troca do uso normal dos arquivos do usuário; o resgate geralmente é pago de forma não rastreável, como bitcoins ).
- Recursos relacionados a recursos relacionados à exploração que constituem malware tentando explorar vulnerabilidades, falhas ou defeitos conhecidos ou não compactados no sistema, ou um ou mais de seus componentes causar instabilidades e comportamentos imprevistos no hardware e no software.
- Recursos relacionados a ladrões de informações, programas maliciosos que coletam informações confidenciais do computador afetado.

Além de detectar comportamentos suspeitos, como chamadas de Application Programming Interface (API)(Interface de Programação de Aplicação), a análise dinâmica também permite reconstituir (limpar) o Operational System (OS) (Sistema Operacional) auditando os malwares promovidos pelo arquivo malicioso no OS, assumindo que o dano não é estatisticamente irreversível. Ressalta-se que a reconstituição do SO, tecnicamente chamada de vacina, é importante porque não basta detectar e eliminar o malware para libertar a vítima de suas ações. Caso nenhuma auditoria seja fornecida pela análise dinâmica, caberá ao observador cibernético monitorar manualmente quaisquer alterações no sistema operacional que alterem o processo lento e estressante.

## 5.2 Classificadoress

O antivírus autoral emprega redes neurais artificiais como classificadores. Para escolher a melhor configuração da arquitetura da rede neural, empregam-se diferentes funções de aprendizado e configurações iniciais para exigir um número maior de cálculos, como multiplicar o número de neurônios na camada intermediária. As arquiteturas de rede neural possuem uma camada de entrada contendo muitos neurônios referentes ao vetor de características extraídas do monitoramento do arquivo CRX em um ambiente controlado. Portanto, os classificadores empregados devem possuir uma camada de entrada contendo 1.098 neurônios. Eles dizem respeito aos recursos de auditoria do arquivo CRX. A camada de saída possui dois neurônios, correspondentes a amostras benignas e de malware.

O trabalho proposto resultou em um antivírus composto por redes neurais de máquinas de aprendizado extremo (Extreme Learning Machines ELMs) para detectar malwares preventivamente. Os ELMs são máquinas de aprendizado baseadas em kernel poderosas e flexíveis, cujas principais características são treinamento rápido e desempenho de classificação robusto (HUANG, 2012). A rede ELM é uma rede de camada oculta única, não recorrente, baseada em um método analítico para estimar os pesos de saída da rede em qualquer inicialização aleatória dos pesos de entrada.

Os ELMs têm sido amplamente aplicados em diversas áreas, como a engenharia biomédica (AZEVEDO, 2015a)(AZEVEDO, 2015b)(AZEVEDO, 2020)(LIMA; SILVA-FILHO; SANTOS, 2014)(LIMA; SILVA-FILHO; SANTOS, 2020)(LIMA; SILVA-FILHO; SANTOS, 2016) (PEREIRA, 2020). As redes ELM podem contribuir muito para o avanço da segurança digital dos dispositivos. O trabalho proposto aplica ELMs na área de segurança da informação especificamente no reconhecimento de padrões de malware.

Matematicamente, na rede neural ELM os atributos de entrada $x_{ti}$ correspondem ao conjunto $\left\{x_{ti} \in \mathbb{R};\ t = 1, ...v\ , i = 1, ...n;\right\}$. Portanto existem $n$ recursos extraídos do aplicativo e vetores de dados de treinamento $v$. A camada oculta $h_j$, composta por $m$ neurônios, é representada pelo conjunto $\left\{h_j \in \mathbb{R}; j \in N^*; j = 1, ..., m\right\}$ O processo de treinamento do ELM é rápido porque é composto de apenas algumas etapas. Inicialmente, os pesos de entrada $w_{ji}$ e bias $bias\ b_{jt}$ são definidos em uma geração aleatória. Dada uma função de ativação $f : \mathbb{R} \rightarrow \mathbb{R}$,, o processo de aprendizagem é dividido em três etapas:

- Geração aleatória do peso $w_{ji}$, correspondente aos pesos entre a entrada e as camadas ocultas, e polarização $bias\ b_{jt}$.
- Calcular a matriz H, que corresponde à saída dos neurônios da camada oculta.
- Calcular a matriz dos pesos de saída $\beta = H^{\dagger}Y$, onde $H^{\dagger}$ é a matriz inversa de Moore-Penrose generalizada da matriz H. A variável Y corresponde à matriz de saídas desejadas, onde $\left\{Y_{tc} \in \mathbb{R};\ t = 1, ...v;\ c = 1, ...\zeta\right\}$. $\zeta$ corresponde às classes (ex. benigno, malware).

O conceito de matriz inversa está relacionado com a matriz identidade. Uma matriz quadrada original $H$ multiplicada por sua inversa $H^{-1}$ é igual à matriz identidade $H.H^{-1} = I$. No entanto, nos casos de matriz dimensional retangular, portanto não quadrada, é gerada uma matriz aproximadamente inversa $H^{\dagger}$. Essa matriz aproximadamente inversa é responsável por polarizar os pesos sinápticos entre os neurônios. A matriz pseudo-inversa repele os pesos sinápticos da fronteira de decisão em direção aos extremos (pólos) da diagonal secundária.

Em termos matemáticos, a matriz pseudo-inversa $H^{\dagger}$ usa a decomposição de valor singular $H = U\Sigma V^{*}$, onde $U$ é uma matriz unitária real ou complexa $n{\times}n$ e $n$ é o neurônio de entrada total. $\Sigma$ é uma matriz diagonal retangular $n{\times}\sigma$ com números reais não negativos na diagonal principal, $n$ é o total neurônio de entrada e $\sigma$ é o total de vetores de dados de treinamento. V* (o conjugado transposto de V) é uma matriz unitária real ou complexa $\sigma{\times}\sigma$. As entradas diagonais $\Sigma_{i,t}$ de $\Sigma$ são denominadas valores singulares de $H$. As colunas $n$ de U e as colunas $\sigma$ de V são os vetores singulares à esquerda e os vetores singulares à direita de $H$, respectivamente. O pseudo-inverso de $H$ é então igual a $H^{\dagger} = V\Sigma^{-1}U^{*}$.

A saída dos neurônios da camada oculta, correspondente à matriz H, é calculada pelo kernel $\varphi$, entradas do conjunto de dados e pesos entre a entrada e as camadas ocultas mostradas na Eq. (5.1). A matriz das saídas obtidas $Y$ e os pesos de saída $\beta$ são definidos em Eq. (5.2) e Eq. (5.3), respectivamente.

$$H_{jt} = \begin{bmatrix} \varphi(1,1) & \varphi(1,2) & \cdots & \varphi(1,v) \\ \varphi(2,1) & \varphi(2,2) & \cdots & \varphi(2,v) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi(m,1) & \varphi(m,2) & \cdots & \varphi(m,v) \end{bmatrix}_{m\times v} \tag{5.1}$$

$$\beta_{jc} = \begin{bmatrix} \beta_1^1 & \cdots & \beta_\zeta^1 \\ \beta_1^2 & \cdots & \beta_\zeta^2 \\ \vdots & \ddots & \vdots \\ \beta_1^m & \cdots & \beta_\zeta^m \end{bmatrix}_{m\times\zeta} \tag{5.2}$$

$$Y_{tc} = \begin{bmatrix} Y_1^1 & \cdots & Y_\zeta^1 \\ Y_1^2 & \cdots & Y_\zeta^2 \\ \vdots & \ddots & \vdots \\ Y_1^v & \cdots & Y_\zeta^v \end{bmatrix}_{v\times\zeta} \tag{5.3}$$

O aprendizado da rede ELM é baseado no kernel. O aprendizado dotado de kernel oferece a possibilidade de criação um mapeamento não-linear de dados sem a necessidade de aumentar

o número de parâmetros ajustáveis, como a taxa de aprendizado comumente usada em redes neurais baseadas em retropropagação. A Eq. (5.4) descreve um kernel Sigmoide $\varphi$ de uma rede ELM com os resultados mostrados na Figura 2 (a).

$$\varphi(t,i) = Sigmoide\Big(x_{ti}.w_{ji} + b_{jt}\Big),$$
$$onde\ Sigmoide(\xi) = \frac{1}{1+e^{-\xi}} \tag{5.4}$$

Ao invés de usar kernels convencionais, os kernels autorais são usados para ELMs. Foi empregado morphological ELM (mELM) (ELMs morfológicos), ELMs com núcleos de camada oculta baseados nos operadores morfológicos de processamento de imagens de erosão e dilatação. Kernels são funções matemáticas empregadas como um método para aprender redes neurais. Este método de aprendizagem permite a criação de mapeamento de dados não linear. Assim, não há necessidade de aumentar o número de parâmetros ajustáveis, como na taxa de aprendizado utilizada em redes com propagação para trás. Existem duas operações morfológicas fundamentais, erosão e dilatação. A teoria da morfologia matemática pode ser considerada construtiva, pois todas as operações são construídas com base na erosão e dilatação. Matematicamente, a erosão e a dilatação são caracterizadas de acordo com a Eq. (5.5) e Eq. (5.6), respectivamente:

$$\epsilon_g(f)(u) = \bigcap_{v \in S} f(v) \ \vee \ \overline{g}(u - v) \tag{5.5}$$

$$\delta_g(f)(u) = \bigcup_{v \in S} f(v) \ \wedge \ g(u - v) \tag{5.6}$$

Onde $f : S \to \{0,1\}$ e $g : S \to \{0,1\}$ são imagens normalizadas na forma de uma matriz denominada $S$, onde $S \in \mathbb{N}^2$ formato. Pixel é definido pelo par cartesiano $(u, f(u))$, onde $u$ é a posição relativa ao valor $f(u)$. $v$ é a matriz de $f(u)$, englobada por $g$. Os operadores estão associados à operação máxima, enquanto $u$ e $v$ estão associados à operação mínima. g é o elemento estruturante tanto para a erosão quanto para a dilatação (SANTOS, 2011). $\overline{g}$ é a negação de $g$.

Na Eq. (5.5), o elemento estruturante $\overline{g}$ é inicialmente negado. Então a operação de máximo $\vee$ denotada por $f(v) \vee \overline{g}(u - v)$, onde $f(v)$ se refere à matriz de imagem original $f$ coberta (combinada) por $overlineg.f(v)$ é tecnicamente chamado de região ativa da imagem.

Por fim, o valor $\epsilon_g(f)(u)$, na posição $u$, da imagem erodida recebe o valor mínimo entre os máximos via o operador $\bigcap$. $\epsilon_g(f)(u)$ obtém o valor 0 associado ao preto absoluto. A erosão sobrepõe $\overline{g}$ à imagem original $f$. O objetivo é que áreas semelhantes a $\overline{g}$ expandam (SANTOS,

2011). Ao associar os 1's ao branco absoluto e os 0's ao preto absoluto, a erosão realça as áreas mais escuras e elimina as regiões de alta intensidade (SANTOS, 2011).

A Eq. (5.6) mostra a operação de dilatação morfológica. Devido à precedência matemática, ocorre a operação mínima $\wedge$ denotada por $f(v) \wedge g(u - v)$, onde $f(v)$ se refere à matriz de imagem original $f$ coberta (combinada) por $g$. Portanto, o $\delta_g(f)(u)$ na posição $u$ da imagem expandida recebe o valor máximo entre os mínimos através do operador $\cup$. A dilatação sobrepõe o elemento estruturante $g$ na imagem original $f$. O objetivo é que áreas semelhantes a $g$ se expandam. Ao associar 1's ao branco absoluto e 0's ao preto absoluto, a dilatação aumenta as áreas com tonalidade mais intensa e elimina as regiões escuras (SANTOS, 2011).

Ao se fazer uma analogia entre a operação de processamento de imagem e o kernel de rede neural mELM, a região ativa da imagem $f(v)$ está associada aos atributos de entrada $x_{t1}, x_{t2}, ..., x_{tn}$. O elemento estruturante $g$ está associado aos pesos $w_{ji}$ das conexões sinápticas entre as camadas de entrada e a escondida.

O antivírus autoral utiliza mELMs. Eles são inspirados na morfologia matemática baseada em operadores não lineares de erosão e dilatação. Como indicado pela Eq. (5.5) em relação ao operador de imagem de erosão, o kernel ELM de erosão pode ser definido juntamente com a Eq. (5.7), onde $\left\{i \in \mathbb{N}^*, i = 1, \ldots, n\right\}$, $\left\{j \in \mathbb{N}^*, j = 1, \ldots, m\right\}$, $\left\{t \in \mathbb{N}^*, t = 1, \ldots, v\right\}$. Existem $n$ neurônios na camada de entrada (sem viés), $m$ neurônios na camada oculta e $v$ vetores de dados de treinamento.

$$\varphi_\epsilon(t,i) = \bigcap_{i=1}^{n}(x_{ti} \vee \overline{w}_{ji}) + b_{jt} \tag{5.7}$$

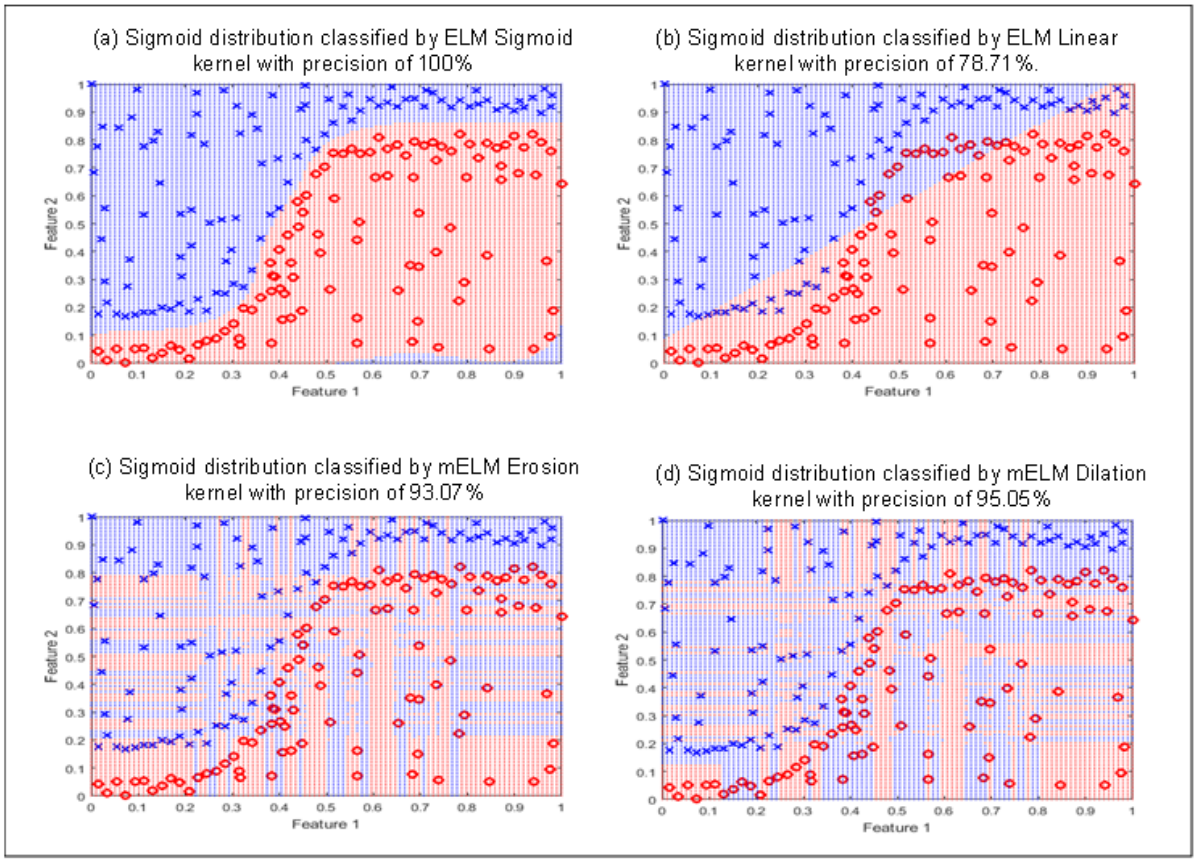
Semelhante ao kernel de erosão, a Eq. (5.8) define o kernel de dilatação inspirado na Eq. (5.6) e refere-se ao operador morfológico de dilatação.

$$\varphi_\delta(t,i) = \bigcup_{i=1}^{n}(x_{ti} \wedge w_{ji}) + b_{jt} \tag{5.8}$$

Para obter um bom desempenho em ELMs, é necessário escolher um kernel que possa otimizar o limite de decisão para o problema apresentado, como visto na Figura 2(a). Um kernel linear obtém ótimos resultados quando usado para resolver um problema linearmente separável. No entanto, quando usado para resolver problemas não linearmente separáveis, como mostra a Figura 2 (b), para uma distribuição sigmóide, não apresenta um desempenho satisfatório. Uma grande capacidade de generalização da rede neural pode depender de uma escolha de kernel bem ajustada. O melhor kernel pode estar subordinado ao problema a ser resolvido.

Como impacto colateral, investigar diferentes kernels é um assunto estressante que envolve validação cruzada combinada com diferentes condições iniciais aleatórias. Entretanto, a

Figura 2 – (a) Desempenho bem-sucedido do kernel compatível com o conjunto de dados. (b) Classificação imprecisa do kernel linear em uma distribuição não linearmente separável. (c - d) Desempenhos bem-sucedidos dos kernels de dilatação e erosão.




Fonte:O autor (2022).

investigação de kernels distintos pode ser necessária; caso contrário, a rede neural é composta por um kernel incompatível para gerar resultados insatisfatórios.

A Figura 2 (c) e a Figura 2 (d) mostram o desempenho da erosão e dilatação do kernel mELM, com acurácias respectivas de 93,07% e 95,05%. É possível perceber ao analisar os números que os mELMs têm a capacidade de mapear com precisão as diferentes distribuições referentes a diferentes problemas.

A eficácia de nossas redes neurais morfológicas se deve à sua capacidade de se adaptar a qualquer tipo de distribuição, uma vez que seu mapeamento não obedece a nenhuma figura geométrica convencional. O mapeamento das fronteiras de decisão é feito por seus dados de treinamento, a própria posição no espaço $n$-dimensional que determinará se aquela região circundante pertence à classe 1 ou à classe 2, $n$ representa o número de neurônios em a camada de entrada. Portanto o kernel mELM trabalhado pode detectar e modelar naturalmente as regiões $n$-dimensionais divididas em diferentes classes usando Morfologia Matemática.

Para provar o embasamento teórico, o antivírus autoral emprega redes neurais morfológi-

cas superficiais em vez de redes convolucionais profundas. Não é a complexidade computacional que tornará a rede neural mais eficiente. A adequação ao problema alvo torna a rede neural eficiente independentemente do número de cálculos. Uma rede linear rasa, por exemplo, pode resolver um problema linearmente separável, bem como uma rede de aprendizado profundo de última geração que consome dias para concluir seu treinamento.

Os kernels morfológicos autorais apresentam uma importante relação com o repositório criado. A justificativa é que a morfologia matemática pode detectar e segmentar os limites dos objetos alvo, preservando as relações dos corpos através do uso da teoria matemática da interseção, união e diferença de conjuntos (SANTOS, 2011). Ao considerar o exemplo contido na Tabela 3, kernels morfológicos autorais podem processar regiões totalmente segregadas, preservando seus limites. Por região, foi denotado uma área contendo valores continuamente congruentes.

# 6 RESULTADOS DAS REDES NEURAIS ELM

Empregamos sete tipos de kernel diferentes para as redes neurais do ELM. No estado da arte, três desses kernels foram descritos por HUANG *et al.* (2012): transformada *wavelet*, *hard limit* e *tribas* (função de base triangular) (HUANG, 2012). Além disso, são empregados núcleos autorais: Dilatação*Fuzzy*, Erosão *Fuzzy*, Dilatação e Erosão.

Os kernels fuzzy-morphological ELM (fmELM) (ELMs *fuzzy-morfológicos*) têm sido bem sucedidos no tratamento de imagens biomédicas, especificamente na detecção e classificação do câncer de mama (AZEVEDO, 2015a)(AZEVEDO, 2015b). Os fmELMs constituem funções lineares, inspiradas na morfologia matemática, com tempo de aprendizado otimizado em relação aos mELMs. Apesar do baixo tempo de aprendizado, os fmELMs não são completamente adequados para distribuições não lineares, como distribuições sigmóides e senoidais.

O kernel wavelet não possui camada oculta. Os cálculos são baseados na transformação dos dados de entrada e podem funcionar de forma semelhante aos kernels contendo arquiteturas com camadas (HUANG, 2012). Uma boa capacidade de generalização desses canais depende de uma escolha ajustada dos parâmetros $(C, \gamma)$ (HUANG, 2012). O parâmetro de custo C refere-se a um ponto de equilíbrio razoável entre a largura da margem do hiperplano e a minimização do erro de classificação em relação ao conjunto de treinamento. O parâmetro do kernel $\gamma$ controla o limite de decisão em função das classes (HUANG, 2012). Não existe um método universal no sentido de escolher os parâmetros $(C, \gamma)$.

No presente trabalho, há uma investigação dos parâmetros $(C, \gamma)$ inspirada no método proposto por HUANG *et al.* (2012), que consiste em treinar sequências crescentes de C e $\gamma$, matematicamente, $2^n$, onde $n = \left\{-24, -10, 0, 10, 25\right\}$ (HUANG, 2012). A hipótese é verificar se esses parâmetros com valores diferentes dos padrões $(C = 1, \gamma = 1)$ geram melhores resultados.

Cada combinação emprega validação cruzada através do método k-fold, onde k=10. O objetivo é que os resultados alcançados não sejam influenciados pelos conjuntos de treinamento e teste. Para isso, o número total de amostras é dividido em dez partes. Na primeira iteração, a primeira parte é o conjunto de teste, enquanto as demais são reservadas para treinamento. Essa rotação ocorre por dez ciclos até que todas as dez peças tenham sido aplicadas à fase de teste.

A acurácia do ELM é a média aritmética da taxa de acertos obtida nos dez ciclos. Conforme mencionado anteriormente, na rede ELM, não há retropropagação de dados. Portanto, o objetivo do método de validação cruzada k-fold não é estabelecer um critério de parada para evitar overfitting (excesso de treinamento), mas verificar se o classificador sofre mudanças abruptas em sua acurácia dependendo dos conjuntos de treinamento e teste. Além disso, cuidados metodológicos devem ser tomados para selecionar aleatoriamente amostras benignas e

de malware para cada dobra. O objetivo é que tendencioso classificadores, em relação a uma determinada classe, não têm seus índices de acertos favorecidos.

A Tabela 5 detalha os resultados obtidos pelas redes neurais ELM com kernel wavelet. Por 45 vezes, um pacote distinto de exemplares benignos (classe majoritária) é apresentado ao único pacote de exemplares malware (classe minoritária). Em cada uma dessas 45 vezes, há a validação cruzada através do método k-fold onde k=10. Então, há 450 (45*10) execuções em cada linha da Tabela 5. Em relação à acurácia na fase de teste, o desempenho médio máximo é de 73,33% na distinção entre casos benignos e de malware com os parâmetros $(C, \gamma) = (2^{10}, 2^{25})$. Na Tabela 5, há apenas as descrições do melhor e do pior caso, nesta ordem, para cada kernel do ELM.

Tabela 4 – Resultados das redes neurais ELM. Os parâmetros $(C, \gamma)$ variam de acordo com o conjunto $\left\{2^{-24}, 2^{-10}, 2^{0}, 2^{10}, 2^{25}\right\}$. São exibidas apenas as melhores e piores acurácias.

| kernel | $(C, \gamma)$ | Acurácia de treinamento (%) | Acurácia de teste (%) | Tempo de treino (seg.) | Tempo de teste (seg.) |
|---|---|---|---|---|---|
| Wavelets | $(2^{10}, 2^{25})$ | **93,96 ± 1,77** | **73,33 ± 19,62** | 0,03 ± 0,01 | 0,00 ± 0,01 |
| | $(2^{0}, 2^{-24})$ | 100,00 ± 0,00 | 50,00 ± 0,00 | 0,04 ± 0,02 | 0,00 ± 0,01 |

Fonte: O autor (2022).

As redes de limite rígido, *tribas* (função de base triangular), fuzzy-Dilation, e fuzzy-erosion, Dilation e erosão empregam arquiteturas de camada oculta. Neste ponto, há uma investigação sobre o número de neurônios na camada oculta desses kernels. A hipótese é verificar se arquiteturas que exigem um volume maior de cálculos, como multiplicar o número de neurônios na camada oculta, podem produzir melhores taxas de acurácia em comparação com arquiteturas que demandam menos cálculos. Duas arquiteturas são avaliadas; eles empregam 100 e 500 neurônios em suas respectivas camadas ocultas. Essas arquiteturas possuem um histórico de excelente acurácia na aplicação de redes ELM na área de engenharia biomédica (LIMA; SILVA; LUZ, 2021).

A Tabela 6 detalha os resultados obtidos pelas redes neurais ELM com os kernels hard limit, *tribas* (função de base triangular), fuzzy-dilatação e fuzzy-erosão, diltação e erosão. Cada linha da Tabela 6 contém 450 execuções, assim como na Tabela 5. Em relação à acurácia, o desempenho médio máximo é de 99,99% com desvio padrão de 0,01% através do kernel Dilation dotado de 500 neurônios em sua camada oculta.

Tabela 5 – Resultado da redes neurais ELM. O número de neurônios na camada oculta varia de acordo com 100, 500.

| kernel | neurônios | Acurácia de treinamento (%) | Acurácia de teste (%) | Tempo de treino (seg.) | Tempo de teste (seg.) |
|---|---|---|---|---|---|
| Hard limit | 100 | 50,00 ± 0,00 | 50,00 ± 0,00 | 0,06 ± 0,02 | 0,00 ± 0,00 |
| | 500 | 50,00 ± 0,00 | 50,00 ± 0,00 | 0,05 ± 0,02 | 0,00 ± 0,00 |
| Tribas | 100 | 50,00 ± 0,00 | 50,00 ± 0,00 | 0,05 ± 0,02 | 0,00 ± 0,00 |
| | 500 | 50,00 ± 0,00 | 50,00 ± 0,00 | 0,05 ± 0,02 | 0,00 ± 0,00 |
| Seno | 500 | 100,00 ± 0,00 | 75,42 ± 19,55 | 0,03 ± 0,01 | 0,00 ± 0,00 |
| | 100 | 100,00 ± 0,00 | 77,50 ± 17,49 | 0,04 ± 0,02 | 0,00 ± 0,00 |
| Dilatação | 100 | **100,00 ± 0,00** | **99,99 ± 0,01** | 0,32 ± 0,08 | 0,02 ± 0,01 |
| | 500 | **100,00 ± 0,00** | **99,99 ± 0,01** | 0,60 ± 0,09 | 0,03 ± 0,01 |
| Erosão | 100 | 100,00 ± 00,00 | 97,50 ± 7,63 | 0,29 ± 0,07 | 0,01 ± 0,01 |
| | 500 | 100,00 ± 00,00 | 97,50 ± 7,63 | 0,51 ± 0,12 | 0,02 ± 0,01 |

Fonte: O autor (2022).

# 7 RESULTADOS EM RELAÇÃO AO ESTADO DA ARTE

Nesta seção, antivírus autorais são comparados com antivírus de última geração. Para evitar comparações injustas, o estágio de extração de recursos é padronizado pelo monitoramento de 1.098 comportamentos que o arquivo CRX suspeito pode realizar quando executado propositalmente. Os antivírus autorais empregam redes neurais morfológicas rasas. A rede neural adotada pelo antivírus autoral consiste de uma camada de entrada com 1.098 neurônios, consistente com o número de atributos de entrada, uma camada oculta de 500 neurônios, uma camada de saída com 2 neurônios, que está de acordo o número de classes, sendo a rede neural do tipo extremo (ELM) e por fim possuindo um kernel de Dilatação.

Em contraste, o antivírus feito por LIMA *et al.* (2021) emprega redes neurais superficiais baseadas em retropropagação. LIMA *et al.* (2021) investigou onze funções distintas de aprendizado para otimizar a precisão de seus antivírus. Para cada função de aprendizagem, LIMA *et al.* (2021) explorou 4 arquiteturas de camada oculta.

O antivírus autoral também é comparado a antivírus baseados em redes neurais profundas. No presente trabalho, os antivírus feitos por SU, *et al.* (2018), VINAYAKUMAR, R. *et al.* (2019), FARUKI, *et al.* (2019), MANIATH, *et al.* (2017), HOU, S. *et al.* (2016), and HARDY, *et al.* (2016), são replicados. O firewall feito por WOZNIAK, M. *et al.* (2015) também é reproduzido. Foram replicados esses trabalhos de aprendizado profundo empregando o conjunto de dados apresentado. As réplicas das obras do estado-da-arte ocorre porque esses antivírus visam diversos dispositivos (Desktop, IoT..). De forma oposta, comparações injustas poderiam ocorrer.

Como dito no Capítulo 4, há 45 pacotes de amostras (22 malware: 990 benginos). A execução de 45 pacotes poderia custar meses visto que os antivírus do estado-da-arte são de Aprendizado Profundo. Devido a essa inviabilidade, foram escolhidos 3 (três) pacotes. Os critérios foram: (i) tempo de treino mais lento, (ii) tempo de treino mais rápido e (iii) tempo de treino médio. Cada réplica de antivírus realiza 30 execuções. Há validação cruzada do método k-fold, onde k = 10, em 3 (três) lotes de dados diferentes.

As Figuras 3 e 4 são representações gráficas dos resultados descritos na Tabela 6. A Figura 3 (a) apresenta os boxplots da etapa de treinamento em relação aos antivírus autorais e métodos de última geração. A melhor média de acerto, resultante do treinamento, é de 99,99% por meio de antivírus autoral. Deep learning feito por MANIATH, *et al.* (2017) e HARDY, et al. (2016) obtiveram uma precisão de treinamento de 100,00% em média. O antivírus feito por LIMA *et al.* (2021) obteve acurácias médias de 69,58% e 100,00% em seus piores e melhores cenários, respectivamente. Esses resultados são obtidos usando as funções de aprendizado "regras de treinamento-aprendizado em lote"e "retropropagação de gradiente conjugado com atualizações de Fletcher-Reeves", respectivamente, com 100 neurônios em suas camadas ocultas.

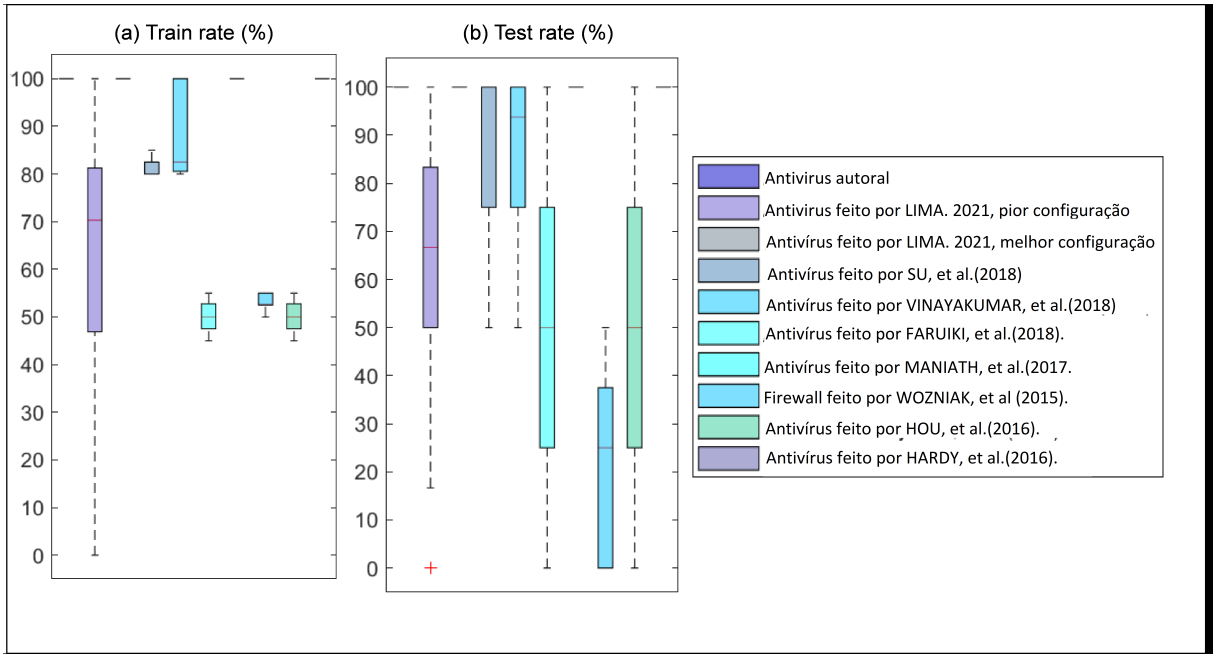



Figura 3 – Boxplots referentes às acurácias do antivírus autoral e ao estado da arte.

Fonte: O autor (2022).

O antivírus autoral obtém desempenho médio de 100,00 ± 0,00% com desvio padrão de 0,00%. Portanto, o antivírus tem a vantagem de não sofrer alterações bruscas devido às condições iniciais (conexões sinápticas e validação cruzada).

A Figura 3 (b) mostra os boxplots para a melhor precisão na fase de teste. Deep learning feito por MANIATH, *et al.* (2017) e HARDY, et al. (2016) obtiveram uma precisão de teste de 99,99% em média. Além disso, o antivírus autoral obtém uma precisão média de 99,99%. O antivírus feito por LIMA *et al.* (2021) atinge uma precisão média de 72,22% e 99,99% em seus piores e melhores cenários, respectivamente. Portanto, corrobora-se que as redes neurais baseadas em retropropagação podem sofrer grandes variações em suas acurácias, dependendo de seus parâmetros de configuração.

Então, a decisão tomada por LIMA, *et al.* (2021) é satisfatória. Este antivírus de última geração explora diferentes funções de aprendizado, gradientes e arquiteturas para otimizar a precisão de suas redes neurais com base na retropropagação de dados. Ressalta-se que sem investigar os parâmetros de redes baseadas em backpropagation, não há garantia de que o antivírus feito por LIMA *et al.* (2021) funcionará sempre em sua melhor configuração.

Uma razão para a falha da maioria dos antivírus baseados em deep learning diz respeito ao uso de um repositório criado a partir de análise dinâmica. Originalmente, os antivírus de última geração empregam análise estática quando o malware geralmente é convertido em uma imagem para servir como atributo de entrada para redes profundas. Esta solução permite o aparecimento de um gradiente vetorial uma vez que as aplicações possuem uma estrutura específica pré-definida. Em contraste, aqui, aplicamos a análise dinâmica. A cadeia de eventos invocada pelos

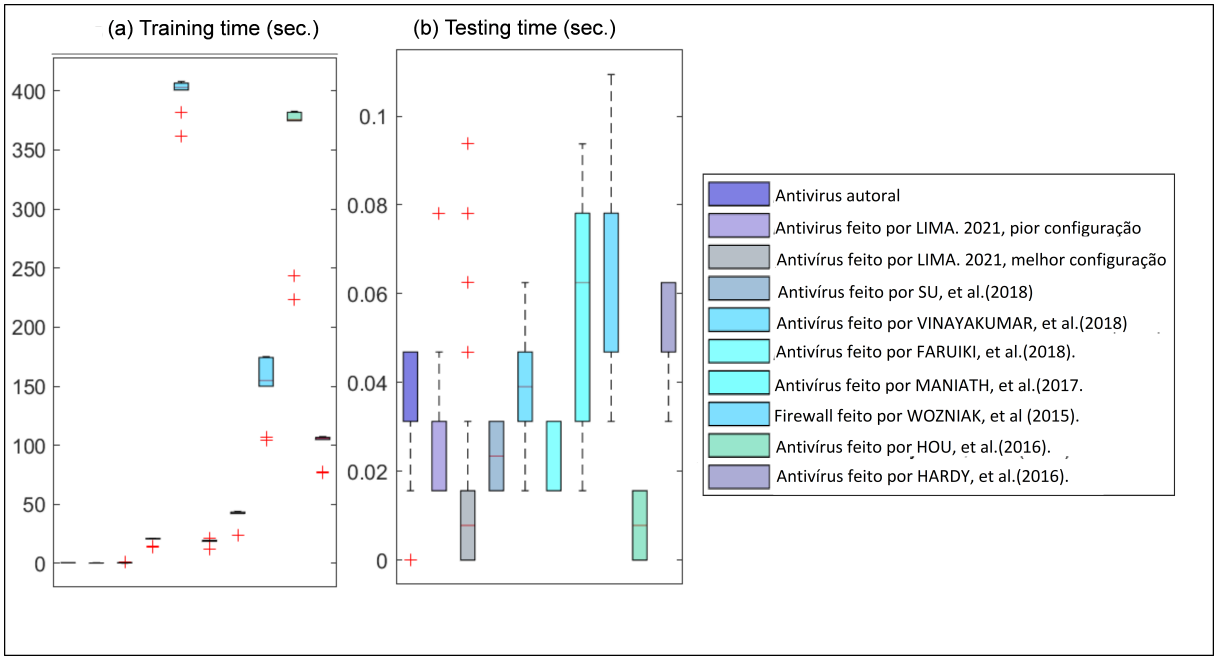



Figura 4 – Boxplots referentes aos tempos de processamento do antivírus autoral e do estado da arte.

Fonte: O autor (2022).

aplicativos suspeitos pode não seguir um gradiente, conforme ilustrado na Tabela 3. Portanto, redes neurais profundas baseadas em convolução de filtro linear podem funcionar fracamente apesar do grande volume de cálculos.

A Figura 4 (a) e a Figura 4 (b) apresentam os boxplots referentes aos tempos gastos nas fases de treinamento e teste, respectivamente. O antivírus autoral consome em média apenas 0,60 segundos para concluir seu treinamento. Em relação ao tempo de treinamento, os antivírus feitos por VINAYAKUMAR, R. *et al.* (2019) e HOU, S. *et al.* (2018) são os mais lentos, pois utilizam uma estrutura recorrente de rede profunda . O antivírus feito por LIMA *et al.* (2021) conclui seu treinamento na ordem de milisegundos porque emprega redes neurais rasas. Em relação ao tempo consumido durante a fase de teste, todas as técnicas consumiram tempos muito próximos sem grandes discrepâncias.

A Tabela 8 mostra as matrizes de confusão das técnicas apresentadas na Tabela 7 em termos percentuais. A matriz de confusão é importante para verificar a qualidade do aprendizado supervisionado. Na Tabela 8, B. e M. são abreviações de benigno e maligno(malware). As classes desejadas estão dispostas na etiqueta vertical, enquanto as classes obtidas estão dispostas na etiqueta horizontal. Em uma matriz de confusão, a diagonal principal é ocupada por casos sempre que a classe obtida coincide com a classe esperada, denominados casos positivos verdadeiros. Então, um bom classificador tem a diagonal principal ocupada por valores altos, e outros elementos têm valores baixos. A Tabela 7 mostra as principais diagonais destacadas em negrito. O antivírus autoral, em fase de teste, classificou erroneamente em média 0,00% dos casos como benignos quando eram casos de malware (falso negativo). Da mesma forma, há uma

classificação média de 0,00% dos casos considerados malware quando são aplicativos benignos (falso-positivo).

Tabela 6 – Comparação entre o antivírus autoral e o estado da arte.

| Técnica | Acurácia de treino (%) | Acurácia de teste (%) | Tempo de treino (seg.) | Tempo de teste (seg.) |
|---|---|---|---|---|
| Antivírus Autoral | **100,00 ± 0,00** | **99,99 ± 0,01** | 0,60 ± 0,09 | 0,03 ± 0,01 |
| Antivirus feito por LIMA, *et al,* (2021), pior conf. | 69,58 ± 22,99 | 72,22 ± 26,08 | 0,40 ± 0,10 | 0,04 ± 0,02 |
| Antivirus feito por LIMA, *et al,* (2021), melhor conf. | **99,00 ± 0,01** | **99,99 ± 0,01** | 0,51 ± 0,19 | 0,02 ± 0,03 |
| Antivirus feito por SU, *et al,* (2018) | 81,81 ± 1,57 | 81,25 ± 15,31 | 17,96 ± 3,94 | 0,02 ± 0,01 |
| Antivirus feito por VINAYAKUMAR, R, *et al,* (2019) | 91,27 ± 8,95 | 87,08 ± 16,89 | 381,54 ± 57,27 | 0,03 ± 0,01 |
| Antivirus feito por MANIATH, S, *et al,* (2017) | **100,00 ± 0,00** | **99,99 ± 0,01** | 37,15 ± 6,27 | 0,06 ± 0,02 |
| *Firewall* feito por WOZNIAK, M *et al,* (2015) | 53,03 ± 1,90 | 21,25 ± 19,74 | 180,33 ± 32,88 | 0,08 ± 0,03 |
| Antivirus feito por HOU, S, *et al,* (2016) | 46,77 ± 13,17 | 50,00 ± 36,41 | 346,90 ± 66,812 | 0,00 ± 0,01 |
| Antivirus feito por HARDY *et al,* (2016) | **100,00 ± 0,00** | **99,99 ± 0,01** | 101,80 ± 11,93 | 0,05 ± 0,01 |
| Antivirus feito por FARUKI *et al,* (2018) | 50,03 ± 3,62 | 51,25 ± 35,26 | 16,46 ± 3,28 | 0,02 ± 0,01 |

Fonte: O autor (2022).

Em relação à Tabela 7, sensibilidade e especificidade referem-se à capacidade do antivírus em identificar malware e aplicativos benignos, respectivamente. O trabalho proposto apresenta a matriz de confusão em termos percentuais para facilitar a interpretação de sensibilidade e especificidade. Em síntese, a sensibilidade e especificidade são apresentadas na matriz de confusão, descrita na Tabela 7. Por exemplo, o antivírus proposto tem uma média de 100,00% tanto em relação à sensibilidade quanto aos verdadeiros positivos. Seguindo o mesmo raciocínio, o antivírus autoral obtém, em média, 100,00% tanto para especificidade quanto para verdadeiros negativos.

Tabela 7 – Matriz de confusão do Antivírus Autoral e Estado da Arte em (%).

| | | Treino | | Teste | |
|---|---|---|---|---|---|
| Técnica | | M. | B. | M. | B. |
| Antivírus Autoral | M, | **100,00 ± 0,00** | 0,00 ± 0,00 | **99,99 ± 0,01** | 0,01 ± 0,01 |
| | B, | 0,00 ± 0,00 | **100,00 ± 0,00** | 0,01 ± 0,01 | **99,99 ± 0,01** |
| Antivirus feito por LIMA, *et al,* (2021), | M, | **68,54 ± 31,68** | 29,38 ± 42,52 | **71,11 ± 34,72** | 26,67 ± 44,98 |
| pior conf. | B, | 31,46 ± 31,68 | **70,63 ± 42,52** | 28,89 ± 34,72 | **73,33 ± 44,98** |
| Antivirus feito por LIMA, *et al,* (2021), | M, | **100,00 ± 0,00** | 0,00 ± 0,00 | **99,99 ± 0,01** | 0,01 ± 0,01 |
| melhor conf. | B, | 0,00 ± 0,00 | **100,00 ± 0,00** | 0,01 ± 0,01 | **99,99 ± 0,01** |
| Antivirus feito por SU, | M, | **100,00 ± 0,00** | 26,73 ± 2,72 | **99,99 ± 0,01** | 38,33 ± 36,20 |
| *et al,* (2018) | B, | 0,00 ± 0,00 | **73,27 ± 2,72** | 0,01 ± 0,01 | **61,67 ± 36,20** |
| Antivirus feito por VINAYAKUMAR, R, | M, | **100,00 ± 0,00** | 12,83 ± 13,16 | **99,99 ± 0,01** | 23,33 ± 32,56 |
| *et al,* (2019) | B, | 0,00 ± 0,00 | **87,17 ± 13,16** | 0,01 ± 0,01 | **76,67 ± 32,56** |
| Antivirus feito pory MANIATH, S | M, | **100,00 ± 0,00** | 0,00 ± 0,00 | **99,99 ± 0,01** | 0,01 ± 0,01 |
| *et al,* (2017) | B, | 0,00 ± 0,00 | **100,00 ± 0,00** | 0,01 ± 0,01 | **99,99 ± 0,01** |
| Deep Learning feito por WOZNIAK, M, | M, | **50,00 ± 50,85** | 50,00 ± 50,85 | **37,50 ± 49,45** | 62,50 ± 49,45 |
| *et al,* (2018) | B, | 50,00 ± 50,85 | **50,00 ± 50,85** | 62,50 ± 49,45 | **37,50 ± 49,45** |
| Antivirus feito por HOU, S, | M, | **6,67 ± 25,37** | 6,67 ± 25,37 | **4,17 ± 20,41** | 4,17 ± 20,41 |
| *et al,* (2018) | B, | 93,33 ± 25,37 | **93,33 ± 25,37** | 95,83 ± 20,41 | **95,83 ± 20,41** |
| Antivirus feito por HARDY, | M, | **100,00 ± 0,00** | 0,00 ± 0,00 | **99,99 ± 0,01** | 0,01 ± 0,01 |
| *et al,* (2016) | B, | 0,00 ± 0,00 | **100,00 ± 0,00** | 0,01 ± 0,01 | **99,99 ± 0,01** |
| Antivirus feito por FARUKI, | M, | **0,00 ± 0,00** | 49,97 ± 3,62 | **0,00 ± 0,00** | 48,75 ± 35,26 |
| *et al,* (2018) | B, | 0,00 ± 0,00 | **50,03 ± 3,62** | 0,00 ± 0,00 | **51,25 ± 35,26** |

Fonte: O autor (2022).

A Tabela 9 mostra os testes de hipóteses paramétricos *t-student* e não paramétricos de *Wilcoxon* entre o antivírus apresentado e o estado da arte. É possível concluir que o antivírus autoral é estatisticamente diferente em comparação aos antivírus de LIMA, *et al.* (2021) (pior conf.), SU, *et al.* (2018), VINAYAKUMAR, *et al.* (2018), HOU, *et al.* (2016) e FARUKI, *et al.* (2019). Assim como o antivírus autoral é estatisticamente distinto em relação ao firewall de WOZNIAK, *et al.* (2015). A explicação é que nessas comparações tanto nos testes paramétricos *t-student* quanto nos não paramétricos de *Wilcoxon*, a hipótese nula foi rejeitada.

Tabela 8 – T-students e Wilcoxon testam as hipóteses do antivírus autoral e do estado da arte.

| | t-students (teste paramétrico) | | Wilcoxon (teste não-paramétrico) | |
|---|---|---|---|---|
| Comparação | Hipóteses | *p*-value | Hipóteses . | *p*-valor |
| Antivírus autoral *vs*<br>Antivirus feito por LIMA, *et al,* (2021), pior conf. | 1 | 1,96554e-16 | 1 | 2,71544e-20 |
| Antivírus autoral *vs*<br>Antivirus feito por LIMA, *et al,* (2021), melhor conf. | 0 | 0 | 0 | 0 |
| Antivírus autoral *vs*<br>Antivírus feito por SU, *et al,* (2018) | 1 | 8,45104e-20 | 1 | 6,13475e-22 |
| Antivírus autoral *vs*<br>Antivírus feito por VINAYAKUMAR, R, *et al,* (2019) | 1 | 9,75754e-11 | 1 | 2,77305e-12 |
| Antivírus autoral *vs*<br>Antivírus feito por MANIATH, S, *et al,* (2017) | 0 | 0 | 0 | 0 |
| Antivírus autoral *vs*<br>Antivírus feito por WOZNIAK, S, *et al,* (2016) | 1 | 4,54717e-57 | 1 | 1,02819e-35 |
| Antivírus autoral *vs*<br>Antivírus feito por HOU, S, *et al,* (2016) | 1 | 1,25427e-22 | 1 | 9,38983e-26 |
| Antivírus autoral *vs*<br>Antivírus feito por HARDY, *et al,* (2016) | 0 | 0 | 0 | 0 |
| Antivírus Autoral *vs*<br>Antivírus feito pela FARUKI, textitet al, (2018) | 1 | 8,38572e-23 | 1 | 9,72978e-26 |

Fonte:O autor (2022).

O antivírus autoral demonstram uma grande vantagem quando comparados aos métodos de última geração. O antivírus atinge um desempenho médio de 100,00% em um tempo médio de treinamento de 0,60 segundos. Essa relação entre precisão percentual e tempo de treinamento em ordem inversa é amplamente empregada na engenharia biomédica (LIMA; SILVA-FILHO; SANTOS, 2016). Admite-se que estabelecer essa relação assume uma função importante na segurança da informação, uma vez que 8 (oito) novos tipos de malware são lançados a cada segundo (INTEL, 2018). Portanto, paradoxalmente, um antivírus recém-lançado pode já estar obsoleto e exigir novo treinamento por meio de uma vulnerabilidade recém-descoberta.

# 8 CONCLUSÃO

A cada ano, milhares de tipos de malware são desenvolvidos em proporções crescentes e contínuas (CUCKOO, 2020). Portanto, é de vital importância que as plataformas de detecção de malware forneçam mecanismos preventivos de vigilância cibernética que atendam às demandas dos clientes. Caso contrário, em cenários em que há falha na identificação do aplicativo malicioso, é provável que dados confidenciais do cliente sejam disponibilizados para pessoas não autorizadas.O AdGuard revelou que mais de 80 milhões de usuários foram induzidos a baixar uma extensão de navegador maliciosa disfarçada de complementos legítimos, como por exemplo adblockers(ADGUARD, 2020).

O presente trabalho investiga os 68 principais antivírus comerciais do mundo quanto à detecção de extensões maliciosas do Google Chrome. A variação na detecção de malware variou de 0% a 63,64%, dependendo do antivírus comercial escolhido. Em média, os antivírus conseguiram detectar 34,95% das extensões maliciosas. Após a análise dos experimentos, foi possível identificar que os antivírus, em média, relataram falsos negativos e foram omitidos em 67,82% e 8,07% dos casos, respectivamente. No presente trabalho, a plataforma VirusTotal foi utilizada para submeter automaticamente malwares a antivírus comerciais.

Vale ressaltar que no VirusTotal não existe a opção de escolher a versão shareware do antivírus. Assim, não foi possível fazer comparações entre funcionalidades comerciais e gratuitas de um mesmo antivírus. É razoável supor que os serviços oferecidos nas versões shareware apresentariam um desempenho significativamente inferior ao das versões completas.

Em média, 41,12% dos antivírus pagos (com todos os recursos) enviados para nossa avaliação não conseguiram detectar nenhuma das extensões maliciosas do Google Chrome. Apesar de movimentar um mercado de bilhões de dólares, os antivírus comerciais não são eficazes em termos de serviços de grande escala e em tempo real. Vale ressaltar que em o estudo apresentado, os malwares analisados são de domínio público, utilizados em atividades maliciosas, e com suas ações catalogadas pelos institutos de pesquisa (VIRUSSHARE, 2022). Mesmo assim, mais de um terço dos antivírus comerciais avaliados não tinham conhecimento da existência do malware investigado.

Para suprir as limitações dos antivírus comerciais, os antivírus baseados em inteligência artificial podem auditar milhares de tipos de malware e aprender, estatisticamente, suas características maliciosas. Portanto, após o aprendizado, os antivírus inteligentes podem identificar e classificar malwares recém-criados de acordo com a comparação entre seus recursos e os catalogados durante sua fase de aprendizado. Assim, não haveria necessidade de esperar um cliente ser contaminado e posteriormente reportar um ataque suspeito como se fosse naquele momento, o antivírus toma alguma ação com relação à descoberta de um novo malware. Os

antivírus inteligentes permitem a detecção preventiva de ameaças virtuais em um ambiente controlado antes que elas cheguem às máquinas dos clientes.

O presente trabalho cria um antivírus capaz de reconhecer preventivamente extensões maliciosas do Google Chrome. Em suma, o antivírus monitora e pondera estatisticamente as ações que o arquivo suspeito pode fazer quando executado no navegador Google Chrome no Windows. Em um ambiente controlado, o antivírus monitora as alterações do registro do Windows, rastreamentos de chamadas feitas por todos os processos gerados por malware, arquivos sendo criados, excluídos e baixados pelo malware enquanto ele está em execução, análise forense de memória e rastreamento de tráfego de rede. Em vez de se ater a eventos individuais, o antivírus é capaz de reconstruir a cadeia de eventos invocada pelo aplicativo auditado. O reconhecimento de padrões, referente às 6.824 ações suspeitas, é feito por redes neurais extremas.

Em vez de kernels convencionais, kernels autorais são empregados para ELMs. A rede ELM tem como principal característica a velocidade de treinamento e previsão de dados quando comparada às redes neurais convencionais. Neste trabalho, empregamos a rede neural morfológica ELM (mELM), uma ELM com kernel de camada oculta, que é inspirada nos operadores morfológicos de processamento de imagens de erosão e dilatação. O kernel Dilation autoral pode distinguir malware CRX de aplicativos benignos em 99,99% dos casos, acompanhados por um tempo de treinamento de 0,60 segundos.

A explicação do sucesso de nossas máquinas de aprendizado morfológico diz respeito à sua capacidade de modelar qualquer decisão limítrofe, uma vez que seu mapeamento não obedece a superfícies geométricas comuns, como elipses e hipérboles empregadas por sistemas de redes neurais clássicos. O mapeamento de decisão limítrofe, realizado pelos kernels morfológicos autorais, utiliza os valores das amostras reservadas para treinamento. Nossa máquina de aprendizado morfológico interpreta a decisão do limite da rede neural como uma imagem $n$-dimensional, onde $n$ é o número de recursos extraídos, incluindo diferentes formas que podem ser descritas usando morfologia matemática. Portanto, os kernels de máquina morfológica lidam naturalmente com o delineamento e modelagem de regiões mapeadas para diferentes classes de qualquer repositório de aprendizado de máquina.

Como trabalho futuro, o antivírus pode ser fornecido como uma extensão do próprio Google Chrome, evitando que extensões maliciosas sejam instaladas. Outra intenção futura é estender o antivírus para outros navegadores além do Google Chrome. Como adversidade, não há amostras de extensões maliciosas de outros navegadores catalogados por institutos de vigilância cibernética. Experimentos preliminares visam criar extensões de direitos autorais maliciosos. A premissa será reproduzir um ataque cibernético inédito (*ZeroDay*) destinado a infectar outros navegadores além do Google Chrome. Portanto, não haverá necessidade de esperar que um usuário seja infectado e, posteriormente, denunciar uma atitude suspeita, e só então o antivírus comercial agirá de forma reativa.

# REFERÊNCIAS

ADGUARD. **80M People Scammed by Chrome Fake Ad Blockers: the Same Old Song**. Disponível em: https://www.https://adguard.com/en/blog/fake-ad-blockers-part-3.html: [s.n.], 2020. Citado na página 40.

AZEVEDO, W. W. *et al.*. **Fuzzy Morphological Extreme Learning Machines to detect and classify masses in mammograms**. doi: https://doi.org/10.1109/FUZZ-IEEE.2015.7337975: [s.n.], 2015. Citado 2 vezes nas páginas 26 e 32.

AZEVEDO, W. W. *et al.*. **Morphological extreme learning machines applied to detect and classify masses in mammograms**. doi: https://doi.org/10.1109/IJCNN.2015.7280774: [s.n.], 2015. Citado 2 vezes nas páginas 26 e 32.

AZEVEDO, W. W. *et al.*. **Morphological Extreme Learning Machines applied to the detection and classification of mammary lesions**. doi: https://doi.org/10.1016/B978-0-12-819295-5.00003-2: [s.n.], 2020. 1-30 p. Citado na página 26.

BASE. **Authorial Dataset: Retrieval for Google Chrome Extension Malware Analysis**. Disponível em: https://github.com/refade/GoogleChromeExtension: [s.n.], 2022. Citado 5 vezes nas páginas 9, 16, 18, 22 e 24.

CIMPANU, C. **Over 500,000 Users Impacted by Four Malicious Chrome Extension**. Disponível em: https://www.bleepingcomputer.com/news/security/over-500-000-users-impacted-by-four-malicious-chrome-extensions/: [s.n.], 2018. Citado na página 12.

CUCKOO. **Análise Automatizada de Malware**. Disponível em: https://cuckoosandbox.org: [s.n.], 2020. Citado 2 vezes nas páginas 23 e 40.

HUANG, G. B. *et al.*. **Extreme Learning Machine for Regression and MultiClass Classification**. doi: https://doi.org/10.1109/TSMCB.2011.2168604: [s.n.], 2012. v. 42(2). 513-519 p. Citado 2 vezes nas páginas 26 e 32.

INTEL. **McAfee Labs**. Disponível em: https://www.mcafee.com/enterprise/en-us/assets/reports/rp-quarterly-threats-mar-2018.pdf: Accessed on Feb 2020, 2018. Citado 2 vezes nas páginas 14 e 39.

LIMA, S. **Limitation of COTS antiviruses: issues, controversies, and problems of COTS antiviruses**. doi: http://dx.doi.org/10.4018/978-1-7998-5728-0.ch020: In: Cruz-Cunha, M.M., Mateus-Coelho, N.R. (eds.) Handbook of Research on Cyber Crime and Information Privacy, vol. 1, 1st edn. IGI Global, Hershey, 2020. Citado na página 16.

LIMA, S.; SILVA-FILHO, A. G.; SANTOS, W. P. **Detection and classification of masses in mammographic images in a multi-kernel approach**. doi: https://doi.org/10.1016/j.cmpb.2016.04.029: [s.n.], 2016. v. 134. 11-29 p. Citado 2 vezes nas páginas 26 e 39.

LIMA, S.; SILVA, H.; LUZ, J. *et al.*. **Artificial intelligence-based antivirus in order to detect malware preventively**. doi: https://doi.org/10.1007/s13748-020-00220-4: [s.n.], 2021. Citado 3 vezes nas páginas 13, 19 e 33.

LIMA, S. M. L.; SILVA-FILHO; SANTOS, W. P. **Morphological Decomposition to Detect and Classify Lesions in Mammograms.In: Wellington Pinheiro dos Santos; Maíra Araújo de Santana; Washington Wagner Azevedo da Silva. (Org.). Understanding a Cancer Diagnosis**. Disponível em: https://novapublishers.com/shop/understanding-a-cancer-diagnosis/: [s.n.], 2020. 27-64 p. Citado na página 26.

LIMA, S. M. L.; SILVA-FILHO, A. G.; SANTOS, W. P. D. **A methodology for classification of lesions in mammographies using Zernike Moments, ELM and SVM Neural Networks in a multi-kernel approach**. doi: https://doi.org/10.1109/SMC.2014.6974041: [s.n.], 2014. Citado na página 26.

MANIATH, S.; ASHOK, A. **Deep Learning LSTM based Ransomware Detection**. doi: https://doi.org/10.1109/RDCAPE.2017.8358312: [s.n.], 2017. Citado na página 20.

MCAFEE. **Why Web-Based Malware Is the Most Serious Threat to Your Business**. Disponível em: https://www.mcafee.com/enterprise/en-us/assets/light-point/white-papers/wp-most-serious-threat-to-business.pdf: [s.n.], 2017. Citado na página 13.

PEREIRA, J. M. S. *et al.*. **Method for Classification of Breast Lesions in Thermographic Images Using ELM Classifiers. In: Wellington Pinheiro dos Santos; Maíra Araújo de Santana; Washington Wagner Azevedo da Silva. (Org.). Understanding a Cancer Diagnosis**. Disponível em: https://novapublishers.com/shop/understanding-a-cancer-diagnosis/: [s.n.], 2020. 117-132 p. Citado na página 26.

PONEMON. **Cost of Data Breach Study**. [S.l.: s.n.], 2015. Citado na página 12.

SANS. **SANS Institute InfoSec Reading Room. Out with The Old, In with The New: Replacing Traditional Antivirus**. Disponível em: https://www.sans.org/reading-room/whitepapers/analyst/old-new-replacing-traditional-antivirus-37377: Accessed on Feb 2020, 2017. Citado na página 16.

SANTOS, W. P. **Mathematical Morphology In Digital Document Analysis and Processing**. [S.l.]: New York: Nova Science, 2011. v. 8. 159-192 p. Citado 3 vezes nas páginas 28, 29 e 31.

SOPHOS. **Sophos Security made simple. Security Threat Report 2014. Smarter, Shadier, Stealthier Malware.** Disponível em: https://www.sophos.com/en-us/medialibrary/pdfs/other/sophos-security-threat-report-2014.pdf: Accessed on Dec. 2020, 2014. Citado na página 13.

SU, J.; VASCONCELLOS D., t. **Lightweight Classification of IoT Malware Based on Image Recognition**. doi: https://doi.org/10.1109/COMPSAC.2018.10315: [s.n.], 2018. Citado na página 20.

VINAYAKUMAR, R.; SOMAN, K. **DeepMalNet: Evaluating shallow and deep networks for static PE malware detection**. doi: https://doi.org/10.1016/j.icte.2018.10.006: [s.n.], 2018. Citado na página 19.

VIRGILLITO, D. **How to spot a malicious browser extension**. Disponível em: https://resources.infosecinstitute.com/topic/how-to-spot-a-malicious-browser-extension/: [s.n.], 2021. Citado na página 12.

VIRUSSHARE. **Banco de dados de arquivos malware**. Disponível em: https://virusshare.com: [s.n.], 2022. Citado 3 vezes nas páginas 13, 21 e 40.