Image-Based Malware Detection

Subjects: Computer Science, Artificial Intelligence

Contributor: Betty Saridou , Isidoros Moulas , Stavros Shiaeles , Basil Papadopoulos

Image conversion of malicious binaries, or binary visualisation, is a relevant approach in the security community. It has exceeded the role of a single-file malware analysis tool and has become a part of Intrusion Detection Systems (IDSs) thanks to the adoption of Convolutional Neural Networks (CNNs).

image-based malware detection | convolutional neural networks | space-filling curves | machine learning

1. Introduction

Image conversion of malicious binaries, or binary visualisation, is a relevant approach in malware detection and analysis. Recent advancements in Deep Learning (DL) and computer vision have allowed security researchers to successfully incorporate image processing techniques in their arsenal [1]. According to this approach, 2D visualisations of malicious and benign files can be used to train Machine Learning (ML)-based classifiers to detect the existence of malware in new entities. In this regard, binary visualisation is fast becoming a key component for Intrusion Detection Systems (IDSs) that wish to overcome the limitations of traditional signature-based techniques. For example, one such limitation is high signature generation sensitivity imposed by even minor differences in code among malware variants [2][3]. Another restriction that is forced by a signature-based detection scheme is the need for constant (and often manual) database updating [2].

Image-based IDSs will also play a vital role in the future as malware detection tools will have to comply with the EU's General Data Protection Regulation's (GDPR) "privacy by design" approach. According to this, malware detection software should avoid any personal data processing, or employ encryption/pseudoanonimisation techniques in any other case [4][5]. For the moment, GDPR, which came into force in 2018, does not impose specific technical directions for applications that stand between security and personal data processing. At the same time, it releases any data that is processed to a state of anonymity from complying with the policy [5]. In that respect, image conversion of personal data (such as network traffic) can act as a form of masking user credentials and facilitate the development of GDPR-compliant security tools.

Image conversion of malware binaries is therefore a promising tool for the security community. Since the first attempts of visualising binaries [6][7] and the first image-based malware classification system [8], there has been a large volume of published studies experimenting with image generation or training classifiers. For example, some studies have been focused on texture analysis [8][9][10], or DL [2][3][11][12], all of which demonstrated positive results towards malware detection.

2. Malware/Benign File Detection

Ref. [13] proposed a system based on binary visualisation with a fuzzy set-based colouring scheme. The images were then passed to several CNN classifiers to detect malicious behaviour. Ref. [14] presented a pattern-based approach for insider threat classification, where they used image-based feature representation to convert log data into images for Wavelet CNN classification. Ref. [15] proposed a Windows malware detection system using CNN and AlexNet learning models. Image conversion of Windows PE files was also the subject of [16], who performed an extensive experiment of a malware detection technique based on RGB images. More specifically, they tested 15 fine-tuned DL models for feature extraction and 12 ML algorithms as the final classifier, from which the winning models were RegNetY320 and SVM. In their research, they also employed data augmentation and transfer learning. Even though [17] did not achieve high accuracy scores when applying CNN classification to their visualisations, they presented an analysis of the interpretability of the malicious patterns within the images. Ref. [18] proposed a visualisable malware detection method based on multi-dimension dynamic behaviours. However, their analysis was carried out on a limited malware dataset. Ref. [19] introduced an ensemble architecture for malware detection that is adaptable to different types of malware. The authors developed a combination of multiple ML algorithms and used an ensemble technique to improve the accuracy of the model. Ref. [20] also published malware detection using deep CNN in combination with image-based representations of malware samples.

3. Malware Detection and Family Classification

Ref. [21] proposed a hybrid framework for image-based malware classification based on space-filling curves. The authors argued that their approach could effectively capture the relationships between binary code and its associated graphical representation. Their approach used traditional ML algorithms to achieve high accuracy and efficiency, outperforming existing malware classification methods. Ref. [2] presented an enhancement for image-based malware classification using ML with low-dimension normalised input images to reduce the dimensionality of the data. However, as they mentioned, the normalisation process may not be suitable for all types of malware. Ref. [22], who had previously used CNN for malware feature extraction in [23], incorporated fuzzy logic in their approach by using convolutional fuzzy neural networks based on feature fusion and the Taguchi method. Ref. [24] presented a malware detection model that is based on Multi-Feature Fusion (MFF) and Histogram of Dynamic Binary Analysis (HDBA). The authors experimentally demonstrated its ability to detect malware variants and its high accuracy in comparison to existing methods with a detection time of 9.63 ms.

To compensate for “the problem of unpredictable truncation” caused by the use of different image widths in malware visualisation, Ref. [25] proposed a malware classification method based on a fusion of Efficient-Net and 1D-CNN. Their method exhibited high accuracy in detecting malware and the ability to classify different malware families. Ref. [26] introduced S-DCNN, an effective DL model which used image representation of binaries and combined transfer and ensemble learning. The model consisted of three CNNs (ResNet50, Xception, and EfficientNet-B4) and a multilayered perceptron as the final classifier. In their work, Ref. [27] presented a triplet neural network (NN) approach to learning similarities between vectors in the latent space. Their method used evolutionary optimisation and an ensemble network to return results by separating entangled representations. The technique demonstrated the segregation of Trojan-based structural features and demonstrated the lower segregation of Simda-typed data in the latent space. Ref. [28], who employed Generative Adversarial Networks (GANs) in their experiment, presented Iron-Dome, a multi-modal malware detection system. Iron-Dome secures IoT networked systems at runtime by classifying executable-based greyscale images.

The authors in [29][30] proposed a new method for malware detection that combined Markov images and transfer learning. The experiments were successful in terms of accuracy and speed but at the same time, they may require more computational resources compared to traditional ML approaches. Ref. [31] proposed the use of visualisation techniques along with the B2IMG and Gabor Filters as image augmentation methods and various CNN architectures to improve malware detection results. Their results were optimal for the B2IMG technique when combined with coloured images. Ref. [32] discussed the use of transfer learning in a EfficientNet3 CNN to classify malware efficiently. The authors utilised pre-trained deep NNs as a base model and fine-tuned them on malware data to achieve high accuracy. According to them, transfer learning proved that it can significantly reduce the training time and computational resources required. Ref. [33] concentrated on reducing the effort needed for malware data labelling. The authors suggested a technique for choosing a subset of data to serve as prototypes for representing the complete dataset and recommended the use of VGG16.

Ref. [34] presented a new approach to malware detection using greyscale binary visualisation and a hyperparameterised CNN. Their method achieved a high accuracy score, while they also provided the code for the training process. Ref. [35] proposed a multilayer DL approach for malware classification in 5G-enabled Industrial Internet of Things (IIoT) systems. The approach used a combination of CNN to detect malware, even in 5G systems, where malware attacks are becoming increasingly common, offering real-time connectivity. Ref. [36] focused on Windows malware detection by exploring the effectiveness and efficiency of the LightGBM algorithm, a gradient-boosting framework that uses tree-based learning algorithms. The authors claim that their approach was effective and fast, both in detecting and classifying malware. In their article, Ref. [37] experimented with various colour models in their visualisation technique to perform malware classification methods. All colour model techniques exhibited equally high performance when they were paired with feature extraction and the SVM model. Ref. [38] presented BinImg2Vec a method that transformed the binary code of malware samples into images and then used the data2vec framework to embed the images in a low-dimensional space. The embedded images were then used as inputs to a CNN for classification.

Ref. [39] proposed MalCNN, a new enhancement for malicious image classification using NNs. Even though the authors claim that their approach outperformed existing methods in terms of accuracy and efficiency, they did not provide a comprehensive comparison with other methods to validate this claim. Additionally, the article did not mention any limitations or potential drawbacks of the proposed method. Ref. [40] presented a static malware classification approach using greyscale image representation and lightweight CNNs that is suitable for IoT environments. Ref. [41] proposed a new method using a combination of automated transmutation and CNNs. However, these articles did not provide a detailed comparison of their proposed methods with other existing methods. Ref. [42] presented a lightweight CNN for image-based malware classification on embedded systems. The authors claim that their approach is effective in classifying malware on embedded systems because of its limited demands on computational resources. Ref. [43] used a Multi-Layer Perceptron (MLP) model. The

authors evaluated their method on a dataset of malware images to show the efficiency of MLP, which was highly hyperparameterised. Ref. [44] attempted a malware classification method based on a lightweight architecture of CNN called MalShuffleNet, which is designed to reduce the computational cost of the model without sacrificing accuracy.

There were also several studies that incorporated the use of GANs in their methodology. Ref. [45] converted malware executables into images, which were then used to train a GAN for family classification. The authors used a variety of techniques for image generation and ML classification. The results favoured the combination of colourmap images and AC-GAN. Ref. [46] used Auxiliary-classifier GAN (AC-GAN) to generate malware images and then evaluate them. The authors found that fake malware images did not impose a threat to adversarial attacks; however, the models were able to successfully classify generated and real malware images. In their study, Ref. [47] used an augmentation model to boost the classification results of malicious images based on their method B1IMG. The CNN-based achieved superior results using RGB photos as opposed to grey images. Ref. [48] also used GANs to enrich samples from the same family and performed classification with CNN. Ref. [49] presented an analysis of the robustness of image-based malware analysis. More specifically, they compared gist descriptors to CNN that were trained directly on malware images to test their malware obfuscation resistance. According to their experiment, they discovered that gist descriptors were more reliable than CNNs.

4. Network Traffic Malware Detection/Classification

Ref. [50] proposed a DL-based model for malware traffic classification using PCAP (packet capture) data. The model converts PCAP data into images and uses image-based NN models to classify the malware traffic. The authors used vision transformers and CNN. Ref. [51] proposed a method for detecting IIoT malware. More specifically, they proposed an edge computing-based malware detection system that identifies malware by sending massive amounts of IIoT traffic data from smart factories to edge servers. By extracting relevant features from the network traffic images, they used DL techniques to identify malicious behaviour. Ref. [52] proposed an ML-based intrusion detection and response system. The system focused on network profiling and trained ML algorithms on historical attack data to detect intrusions in real time. Ref. [53] proposed a method that framed network flows as images and used image recognition algorithms to detect anomalies. The authors used federated learning to train the algorithms and evaluated the performance of their system on DDoS attack data.

References

1. Sahin, M.; Bahtiyar, S. A Survey on Malware Detection with Deep Learning. In Proceedings of the 13th International Conference on Security of Information and Networks, Merkez, Turkey, 4–7 November 2020; pp. 1–6.
2. Son, T.T.; Lee, C.; Le-Minh, H.; Aslam, N.; Dat, V.C. An enhancement for image-based malware classification using machine learning with low dimension normalized input images. J. Inf. Secur. Appl. 2022, 69, 103308.
3. Vasan, D.; Alazab, M.; Wassan, S.; Safaei, B.; Zheng, Q. Image-Based malware classification using ensemble of CNN architectures (IMCEC). Comput. Secur. 2020, 92, 101748.
4. Stupka, V.; Horák, M.; Husák, M. Protection of personal data in security alert sharing platforms. In Proceedings of the 12th International Conference on Availability, Reliability and Security, Reggio Calabria, Italy, 29 August–1 September 2017; pp. 1–8.
5. Pawlicka, A.; Jaroszewska-Choras, D.; Choras, M.; Pawlicki, M. Guidelines for stego/malware detection tools: Achieving GDPR compliance. IEEE Technol. Soc. Mag. 2020, 39, 60–70.
6. Yoo, I. Visualizing windows executable viruses using self-organizing maps. In Proceedings of the 2004 ACM Workshop on Visualization and Data mining For Computer Security, Washington, DC, USA, 29 October 2004; pp. 82–89.
7. Conti, G.; Dean, E.; Sinda, M.; Sangster, B. Visual reverse engineering of binary and data files. In Proceedings of the International Workshop on Visualization for Computer Security; Springer: Berlin/Heidelberg, Germany, 2008; pp. 1–17.
8. Nataraj, L.; Karthikeyan, S.; Jacob, G.; Manjunath, B.S. Malware images: Visualization and automatic classification. In Proceedings of the 8th International Symposium on Visualization for Cyber Security,

Pittsburgh, PA, USA, 20 July 2011; pp. 1–7.

9. Nataraj, L.; Yegneswaran, V.; Porras, P.; Zhang, J. A comparative assessment of malware classification using binary texture analysis and dynamic analysis. In Proceedings of the 4th ACM Workshop on Security and Artificial Intelligence, Chicago, IL, USA, 21 October 2011; pp. 21–30.
10. Nataraj, L.; Manjunath, B. Spam: Signal processing to analyze malware . IEEE Signal Process. Mag. 2016, 33, 105–117.
11. Ni, S.; Qian, Q.; Zhang, R. Malware identification using visualization images and deep learning. Comput. Secur. 2018, 77, 871–885.
12. Le, Q.; Boydell, O.; Mac Namee, B.; Scanlon, M. Deep learning at the shallow end: Malware classification for non-domain experts. Digit. Investig. 2018, 26, S118–S126.
13. Saridou, B.; Rose, J.R.; Shiaeles, S.; Papadopoulos, B. SAGMAD—A Signature Agnostic Malware Detection System Based on Binary Visualisation and Fuzzy Sets. Electronics 2022, 11, 1044.
14. Randive, K.; Mohan, R.; Sivakrishna, A.M. An efficient pattern-based approach for insider threat classification using the image-based feature representation. J. Inf. Secur. Appl. 2023, 73, 103434.
15. Sai Adhinesh Reddy, T.; Varma Vadlamudi, V.Y.; Acharya, S.; Rawat, U.; Bhatnagar, R. Windows Malware Detection Using CNN and AlexNet Learning Models. In Proceedings of the 8th International Conference on Advanced Intelligent Systems and Informatics, Cairo, Egypt, 20–22 November 2022; Springer International Publishing: Berlin/Heidelberg, Germany, 2022; pp. 271–283.
16. Shaukat, K.; Luo, S.; Varadharajan, V. A novel deep learning-based approach for malware detection. Eng. Appl. Artif. Intell. 2023, 122, 106030.
17. Marais, B.; Quertier, T.; Chesneau, C. Malware analysis with artificial intelligence and a particular attention on results interpretability. In Distributed Computing and Artificial Intelligence, Volume 1: 18th International Conference 18; Springer International Publishing: Berlin/Heidelberg, Germany, 2022; pp. 43–55.
18. Ma, Z.; Zhang, Z.; Liu, C.; Hu, T.; Li, H.; Ren, B. Visualizable Malware Detection based on Multi-dimension Dynamic Behaviors. In Proceedings of the 2022 International Conference on Networking and Network Applications (NaNA), Urumqi, China, 3–5 December 2022; pp. 247–252.
19. Mane, D.T.; Kumbharkar, P.B.; Javheri, S.B.; Moorthy, R. An Adaptable Ensemble Architecture for Malware Detection. In International Conference on Innovative Computing and Communications: Proceedings of ICICC; Springer: Singapore, 2022; Volume 3, pp. 647–659.
20. Malani, H.; Bhat, A.; Palriwala, S.; Aditya, J.; Chaturvedi, A. A Unique Approach to Malware Detection Using Deep Convolutional Neural Networks. In Proceedings of the 2022 4th International Conference on Electrical, Control and Instrumentation Engineering (ICECIE), KualaLumpur, Malaysia, 26 November 2022; pp. 1–6.
21. O'Shaughnessy, S.; Sheridan, S. Image-based malware classification hybrid framework based on space-filling curves. Comput. Secur. 2022, 116, 102660.
22. Lin, C.J.; Huang, M.S.; Lee, C.L. Malware Classification Using Convolutional Fuzzy Neural Networks Based on Feature Fusion and the Taguchi Method. Appl. Sci. 2022, 12, 12937.
23. Lin, C.J.; Lin, X.Y.; Jhang, J.Y. Malware classification using a Taguchi-based deep learning network. Sens. Mater 2022, 34, 3569–3580.
24. Wang, S.; Wang, J.; Song, Y.; Li, S.; Huang, W. Malware Variants Detection Model Based on MFF–HDBA. Appl. Sci. 2022, 12, 9593.
25. Chong, X.; Gao, Y.; Zhang, R.; Liu, J.; Huang, X.; Zhao, J. Classification of Malware Families Based on Efficient-Net and 1D-CNN Fusion. Electronics 2022, 11, 3064.
26. Parihar, A.S.; Kumar, S.; Khosla, S. S-DCNN: Stacked deep convolutional neural networks for malware classification. Multimed. Tools Appl. 2022, 81, 30997–31015.
27. Park, K.W.; Bu, S.J.; Cho, S.B. Evolutionary Triplet Network of Learning Disentangled Malware Space for Malware Classification. In Proceedings of the Hybrid Artificial Intelligent Systems: 17th International

Conference, HAIS 2022, Salamanca, Spain, 5–7 September 2022; Springer International Publishing: Berlin/Heidelberg, Germany, 2022; pp. 311–322.

28. Shukla, S.; Dhavlle, A.; PD, S.M.; Homayoun, H.; Rafatirad, S. Iron-Dome: Securing IoT Networked Systems at Runtime by Network and Device Characteristics to Confine Malware Epidemics. In Proceedings of the 2022 IEEE 40th International Conference on Computer Design (ICCD), Olympic Valley, CA, USA, 23–26 October 2022; pp. 259–262.
29. Kwan, L.M. Markov Image with Transfer Learning for Malware Detection and Classification. In Proceedings of the TENCON 2022—2022 IEEE Region 10 Conference (TENCON), Hong Kong, China, 1–4 November 2022; pp. 1–6.
30. Kiger, J.; Ho, S.S.; Heydari, V. Malware Binary Image Classification Using Convolutional Neural Networks. In Proceedings of the International Conference on Cyber Warfare and Security, Islamabad, Pakistan, 7–8 December 2022; Volume 17, pp. 469–478.
31. Dharmalaksana, P.S.; Mantoro, T.; Khakim, L.; Nurseno, M. Improved Malware Detection Results using Visualization-Based Detection Techniques ant Convolutional Neural Network. In Proceedings of the 2022 IEEE 8th International Conference on Computing, Engineering and Design (ICCED), Sukabumi, Indonesia, 28–29 July 2022; pp. 1–5.
32. AlGarni, M.D.; AlRoobaea, R.; Almotiri, J.; Ullah, S.S.; Hussain, S.; Umar, F. An efficient convolutional neural network with transfer learning for malware classification. Wirel. Commun. Mob. Comput. 2022, 2022, 4841741.
33. Cher, G.; Liu, S. Reducing Malware labeling Efforts Through Efficient Prototype Selection. In Proceedings of the 2022 26th International Conference on Engineering of Complex Computer Systems (ICECCS), Hiroshima, Japan, 26–30 March 2022; pp. 17–22.
34. Omar, M. New Approach to Malware Detection Using Optimized Convolutional Neural Network. In Proceedings of the Machine Learning for Cybersecurity: Innovative Deep Learning Solutions; Springer International Publishing: Berlin/Heidelberg, Germany, 2022; pp. 13–35.
35. Ahmed, I.; Anisetti, M.; Ahmad, A.; Jeon, G. A Multilayer Deep Learning Approach for Malware Classification in 5G-Enabled IIoT. IEEE Trans. Ind. Inform. 2022, 19, 1495–1503.
36. Onoja, M.; Aimufua, G.; Jegede, A.; Oyedele, A.; Mazadu, J.; Olibodum, K. Exploring the Effectiveness and Efficiency of LightGBM Algorithm for Windows Malware Detection. Available online: https://www.researchgate.net/profile/Abayomi-Jegede/publication/366167472_2022_5th_Information_Technology_for_Education_and_Development_ITED/links/63945b6311e9f0 5th-Information-Technology-for-Education-and-Development-ITED.pdf (accessed on 27 January 2023).
37. Chauhan, D.; Singh, H.; Hooda, H.; Gupta, R. Classification of malware using visualization techniques. In International Conference on Innovative Computing and Communications: Proceedings of ICICC; Springer: Singapore, 2022; Volume 3, pp. 739–750.
38. Sern, L.J.; Keng, T.K.; Fu, C.Z. BinImg2Vec: Augmenting Malware Binary Image Classification with Data2Vec. In Proceedings of the 2022 1st International Conference on AI in Cybersecurity (ICAIC), Victoria, TX, USA, 24–26 May 2022; pp. 1–6.
39. Kavitha, P.M.; Muruganantham, B. Mal_CNN: An Enhancement for Malicious Image Classification Based on Neural Network. Cybern. Syst. 2022, 1–14.
40. Belguendouz, H.; Guerid, H.; Kaddour, M. Static Classification of IoT Malware using Grayscale Image Representation and Lightweight Convolutional Neural Networks. In Proceedings of the 2022 5th International Conference on Advanced Communication Technologies and Networking (CommNet), Marrakech, Morocco, 12–14 December 2022; pp. 1–8.
41. Agarwal, R.; Patel, S.; Katiyar, S.; Nailwal, S. Malware classification using automated transmutation and CNN. In Advanced Computing and Intelligent Technologies: Proceedings of ICACIT; Springer: Singapore, 2022; pp. 73–81.
42. Fathurrahman, A.; Bejo, A.; Ardiyanto, I. Lightweight Convolution Neural Network for Image-Based Malware Classification on Embedded Systems. In Proceedings of the 2021 International Seminar on Machine

Learning, Optimization, and Data Science (ISMODE), Jakarta, Indonesia, 29–30 January 2022; pp. 12–16.

43. Ben Abdel Ouahab, I.; Elaachak, L.; Bouhorma, M. Image-Based Malware Classification Using Multi-layer Perceptron. In Networking, Intelligent Systems and Security: Proceedings of NISS; Springer: Singapore, 2022; pp. 453–464.
44. Qiu, L.; Wang, S.; Wang, J.; Wang, Y.; Huang, W. Malware Classification based on a Light-weight Architecture of CNN: MalShuffleNet. In Proceedings of the 2022 3rd International Conference on Computer Vision, Image and Deep Learning & International Conference on Computer Engineering and Applications (CVIDL & ICCEA), Changchun, China, 20–22 May 2022; pp. 1047–1050.
45. Nguyen, H.; Di Troia, F.; Ishigaki, G.; Stamp, M. Generative adversarial networks and image-based malware classification. J. Comput. Virol. Hacking Tech. 2023, 1–17.
46. Nagaraju, R.; Stamp, M. Auxiliary-classifier GAN for malware analysis. In Artificial Intelligence for Cybersecurity; Springer: Berlin/Heidelberg, Germany, 2022; pp. 27–68.
47. Tekerek, A.; Yapici, M.M. A novel malware classification and augmentation model based on convolutional neural network. Comput. Secur. 2022, 112, 102515.
48. Kuo, W.C.; Chen, Y.T.; Huang, Y.C.; Wang, C.C. Malware Detection Based on Image Conversion. In Proceedings of the 2021 International Conference on Security and Information Technologies with AI, Internet Computing and Big-data Applications, Taichung, Taiwan, 18–20 November 2021; Springer International Publishing: Berlin/Heidelberg, Germany, 2022; pp. 180–190.
49. Tran, K.; Di Troia, F.; Stamp, M. Robustness of Image-Based Malware Analysis. In Proceedings of the Silicon Valley Cybersecurity Conference: Third Conference, SVCC 2022, Virtual Event, 17–19 August 2022; Revised Selected Papers. Springer: Berlin/Heidelberg, Germany, 2023; pp. 3–21.
50. Agrafiotis, G.; Makri, E.; Flionis, I.; Lalas, A.; Votis, K.; Tzovaras, D. Image-based Neural Network Models for Malware Traffic Classification using PCAP to Picture Conversion. In Proceedings of the 17th International Conference on Availability, Reliability and Security, Vienna, Austria, 23–26 August 2022; pp. 1–7.
51. Kim, H.M.; Lee, K.H. IIoT Malware Detection Using Edge Computing and Deep Learning for Cybersecurity in Smart Factories. Appl. Sci. 2022, 12, 7679.
52. Rose, J.R.; Swann, M.; Grammatikakis, K.P.; Koufos, I.; Bendiab, G.; Shiaeles, S.; Kolokotronis, N. IDERES: Intrusion detection and response system using machine learning and attack graphs. J. Syst. Archit. 2022, 131, 102722.
53. Toldinas, J.; Venčkauskas, A.; Liutkevičius, A.; Morkevičius, N. Framing Network Flow for Anomaly Detection Using Image Recognition and Federated Learning. Electronics 2022, 11, 3138.