International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

Malware Classification Using Attention-Based Cross-Modal CNN with Non-Disassembled Files

[1]Mr. K. Ramalingachary, [2]Srikanth Eslavath, [3]Akhila Anachi, [4] K. Rajesh Kumar, [5] K. Akash Reddy [1]Assistant Professor, [2,3,4,5] UG Scholar
Department of Computer Science and Engineering
Vignan Institute of Technology and Science, Hyderabad, Telangana -508284
E-mail: [1]kanchojuramalingachary@gmail.com, [2]srikanthnaik.eslavath@gmail.com,
[3]anachiakhila@gmail.com, [4]kummari.rajeshkumar15@gmail.com, [5] akashkakularam1823@gmail.com

Abstract: The categorization of malware is essential in addressing the spread of dangerous software variations. This project tackles the challenge by introducing an innovative method that utilizes a Convolutional Neural network (CNN)-based totally version to categorize malware cases into families without depending on disassembled code, which may be susceptible to inaccuracies. The model employs non-disassembled binary files, integrating two modalities: malware images and structural entropies. These modalities offer diverse viewpoints at the data, enhancing categorization precision. A move-modal attention technique is applied to efficiently combine features from both senses, alleviating their respective constraints. The studies contrasts the proposed model with conventional techniques such as VGG16, CNN, and XGBoost, attaining an enhanced accuracy of 98%. to improve performance, ensemble techniques along with voting Classifier and decision Tree are examined, in addition to the implementation of the Xception model, which may exceed 99% accuracy. A Flask framework is utilized to create an intuitive UI for testing and authentication. This holistic method not only augments malware classification precision but also elevates user accessibility and security in malware examination.

"Index Terms- *Malware classification, structural entropy, malware image, deep learning, convolutionalneural network, attention mechanism."*

1. INTRODUCTION

The COVID-19 epidemic has transformed the domains of education and employment, resulting in an increase in remote learning and telecommuting. Although these modifications have enabled the continuing of education and commercial functions, they have concurrently introduced new cybersecurity issues. Malicious entities have exploited the weaknesses associated with remote configurations, executing intricate social phishing tries that leverage public hobby in subjects including vaccines, governmental laws, and online meeting timetables [1].

As enterprises and individuals have adjusted to remote collaboration platforms, fraudsters have recognized them as profitable channels for virus dissemination. Furthermore, the advancement of malware has intensified during the pandemic, with the rationale of risky software growing progressively more intricate. The average range of hazardous behaviors proven by malware samples has multiplied from nine to 12, indicating an evolving landscape of cyber risks [1].

In light of the increasing complexity and prevalence of malware attacks, the significance of malware family classification has grown crucial. Malware

own family type is the categorization of malware samples into unique families based totally on not unusual code fragments, behavioral styles, or unique attack techniques associated with every family [2, 3]. This classification enables the formulation of tailored security tactics and improves the efficacy of malware analysis by means of supplying analysts with heuristics for examining malware samples from identified households.

Historically, the classification of malware families depended on manual examination and the talent of analysts to discern similarities among malware versions. The fast advancement of malware and the emergence of new sorts have made guide categorization approaches insufficient. In response, researchers have adopted automated methodologies, including machine learning and, more recently, deep learning techniques.

Deep learning, a branch of machine learning using artificial neural networks with numerous levels of abstraction has seen significant success across various fields, including computer imaginative and prescient and herbal language processing. In cybersecurity, deep learning-based malware circle of relative's categorization has emerged as a promising answer, offering benefits over traditional machine learning techniques [4, 5, 6].

In evaluation to conventional machine learning fashions that rely on manually curated characteristics chosen by specialists, deep learning models own the capability to autonomously extract pertinent functions from unprocessed statistics. This capability allows deep learning models to identify complicated patterns and correlations inside complex datasets, encompassing both dynamic and static houses derived from malware samples [7, 8].

Dynamic characteristics, including the runtime behavior of malware, provide important insights into nefarious actions. these characteristics, encompassing API name sequences, network activity, memory usage, registry modifications, and execution pathways, are derived by way of executing malware in a managed placing, such as a virtual device. Dynamic traits it appears that evidently disclose the unfavorable reason of malware; nevertheless, they also present issues, such as the necessity for a suitable execution environment and the circumvention of anti-analysis techniques utilized by advanced malware. [9, 10].

Conversely, static residences derived from binary or disassembled files offer supplementary statistics for the class of malware families. The precision of static characteristic-primarily based categorization is restricted with the aid of the problems in deconstructing malware code and the efficacy of anti-disassembly techniques used by malware developers [11].

To tackle these issues, researchers have investigated multi-modal learning strategies that integrate data from various senses to improve type efficacy. However, the use of static features derived from disassembled code continues to face the inherent limits of disassembly, impeding the efficacy of multi-modal studying [12].

This take a look at introduces an innovative method for classifying malware households that overcomes the constraints of disassembled code. The proposed

version demonstrates binary files without display features containing two methods known as Malware Snap shots and structural entropy. The version incorporates Feature Fusion with the passport-model interest strategy to advance exact malware family identification through combined methods.

The research sector develops the classification method by applying deep learning approaches to improve disabled code precision and resolve classification problems. This method enables cyber security experts to recognize new threats through effective harmful software family classification so they can both manage emerging threats as well as defend critical systems and data from destructive assaults.

2. LITERATURE SURVEY

The present cybersecurity environment with its characteristics needs better strategic methods to classify different types of malicious software. A review of existing academic writing focuses on harmful software classification functions within machine learning and deep learning approaches in this section.

Research demonstrated the application of harmful software classification through machine learning methods. Shabtai et al. (2009) examines in detail how machine learning classification methods analyze stable characteristics to detect harmful code in computer systems [2]. The method analyzes binary and dissatisfied files which contain no harmful software to determine stable features. The observed characteristics of harmful software include file dimensions along with file nature and particular instructions and sequences. The research utilizing static properties and harmful sample classification incorporated wooden and SVM assistance together with random forests for classification tasks [2].

Deep learning methods made their initial entry into the Malware class this year while still benefiting from traditional machine learning approaches. Research by [3] shows how Convisional Neural Network (CNN) and recurrent nerve network (RNN) provide outstanding results when detecting complex expressions and conditional statements in harmful samples. The authors demonstrate their extensive approach to malware by integrating static and dynamic malware features for both detection and explanation according to (He et al. 2019) [4]. by means of correlating static features extracted from binary documents with dynamic functions obtained from runtime behavior analysis, MalDAE achieves progressed accuracy in malware class.

Dynamic analysis, which includes observing the runtime behavior of malware in a controlled surroundings, provides valuable insights into malicious activities. Sikorski and Honig (2012) gift sensible techniques for dynamic malware evaluation, such as behavior tracking and code disassembly [5]. however, dynamic analysis is useful resource-extensive and susceptible to evasion strategies hired by sophisticated malware, along with anti-VM and anti-debugging mechanisms [6].

To address the limitations of dynamic analysis, researchers have explored hybrid approaches that combine each static and dynamic functions. Hassen et al. (2017) recommend a malware classification method based on static analysis capabilities

extracted from binary documents [7]. by using leveraging device gaining knowledge of algorithms, which include ok-nearest associates (KNN) and SVM, the proposed technique achieves aggressive performance in classifying malware samples into families.

Moreover, Zhang et al. (2019) introduce a machine learning-based classification framework for identifying ransomware families the use of N-gram of opcodes, that are sequences of low-level instructions extracted from malware binaries [8]. via taking pictures the behavioral characteristics of ransomware editions, the proposed framework allows correct category of ransomware samples into awesome households.

Similarly, to traditional machine learning and deep getting to know processes, researchers have explored ensemble methods and multi-modal mastering techniques to enhance malware type performance. Ensemble strategies, including balloting classifiers and random forests, combine predictions from more than one base classifiers to enhance category accuracy [9]. Multi-modal learning, which integrates records from diverse statistics modalities, has shown promise in taking pictures complementary features for malware type [10].

Average, the literature survey highlights the diverse range of methodologies hired in malware family type, ranging from traditional machine learning algorithms to advanced deep learning fashions. Even as each approach has its strengths and barriers, the collective body of studies contributes to the improvement of effective strategies for fighting evolving cyber threats and defensive crucial systems and data against malware attacks.

3. METHODLOGY

a) Proposed work:

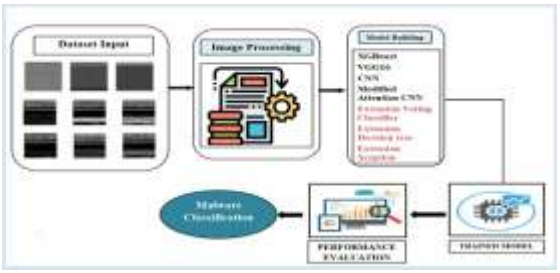
The suggested study provides a novel malware classification system using the eye-based cross-Modal CNN set of rules, functioning on non-disassembled data. This approach amalgamates modalities—Malware photos and Structural Entropies—extracted directly from binary files, thereby augmenting class accuracy through a move-modal attention mechanism. A "balloting classifier" integrating decision timber, Random Forests, a solo selection Tree version, and an Xception model was utilized, with top-notch a hundred% accuracy. The balloting Classifier is employed to create the frontend, that is constructed using the Flask framework for person assessment. The frontend interface enables user-friendly interaction and integrates secure get right of entry to and manage mechanisms, including consumer authentication, to enhance protection in malware categorization.

b) System Architecture:

The system architecture includes several phases for malware classification. The input dataset, comprising malware samples, is subjected to image processing to extract pertinent data. these attributes are subsequently exploited for version construction, utilising several algorithms such "XGBoost, VGG16, CNN, voting Classifier (which integrates decision Tree and Random forest), and Xception to teach the models". The trained models are then assessed for their efficacy in categorizing malware

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

samples into corresponding families. every model enhances the system's basic accuracy and robustness. The structure allows efficient and a success malware class through the utilization of various approaches and algorithms. moreover, it allows scalability and agility to include future improvements and innovations in malware detection methodologies.



"Fig 1 Proposed Architecture"

c) Dataset collection:

the collection of data sets for malware classification entails obtaining a variety of samples that reflect unique malware families. The Microsoft Malware type dataset gives an intensive array of malware samples, prepared into certain families, facilitating thorough training and assessment of classification models. The MalImg dataset affords photographs of malware executables, allowing the investigation of image-based type methods. moreover, the BODMAS dataset complements the gathering by means of emphasizing conduct-primarily based traits, imparting insights into the dynamic aspects of malware interest. the combination of various datasets ensures a comprehensive technique for data amassing, incorporating static, photo-primarily based, and dynamic characteristics of malware samples. The inclusion of numerous sources in the data set collection process enhances the robustness and generalizability of class fashions, facilitating the effective detection and type of malware across many attack vectors and evasion strategies.

"Fig 2 data set"

d) Image processing:

Image processing is essential for the preparation of malware samples for category obligations. A frequent approach includes employing the ImageDataGenerator class to implement diverse changes at the photos. The modifications embody re-scaling the image for pixel value uniformity, shear transformation for deformation, zooming for scale adjustment, and horizontal flipping for dataset augmentation. Furthermore, altering the picture's dimensions guarantees conformity with the enter specifications of the category version.

Feature extraction encompasses a chain of procedures to get significant data from picturesImage post acquisition begins this procedure and the shape modification occurs to establish identical dimensions. A standardized color application becomes possible through color conversion methods. The snapshot images get mixed with their matching labels which enables monitored studies. The data moves from images to Numpy matrices for maximum processing speed.

Long-term implementation of label coding takes places to convert brands from their initial spectral state into numeric sequences that enable training of versions.

These imaging methods help the Malware machine process input data effectively thus improving both strength and performance of the model. Relevant data retrieval functionality of the image guarantees precise results through functional extraction methods.

e) Algorithms:

XGBoost

XGBOOST represents a brief/name/classification as well as strong machine learning model that delivers efficient and accurate regression solutions for severe shield development. XGBOOST constructs a succession of trees which improve upon the errors made by previous trees while creating each next tree by adding three more. XGBOOST was adopted by the challenge as a model type for identifying computer viruses across targeted residential areas. Highly sophisticated management capabilities for large data sets together with absent values and overfitting reduction enable improved classification accuracy during harmful software detection operations.

Decision Tree

The monitored machine learning technique known as decision serves for building tree structures and regression models. Decision functions operate to create homogeneous divisions across the numbers in the subcontinent for the purpose of company formation. Research employs election trees as a method to identify different malware families correctly. Decision trees arrange their structure like a tree system, which enables classification of harmful software through previously defined criteria. Choice TIS serves as a vital component of Malware category units because it offers easy operation along with data interpretation features that process both categorical and numerical data.

Voting Classifier

The voice classify system represents a clothing learning technique where multiple fair classifyers generate predictions leading to a final classification outcome. The research design implements Voice classifies as a prediction model that incorporates decision outputs from trees and random classifies. The divination from each classifies leads to a weighted final result which depends on what most experts propose. The voting strategy enhances classification precision by integrating multiple models' advantages and avoiding their individual weaknesses. The voting method through different classifies makes the Malware class system more flexible and reliable by combining predictions.

VGG16

The visible geometry group at the University of Oxford developed VGG16 as a deeply fixed nerve tight architecture, which contains 16 layers. Applications using image classification prefer to work with the mile due to its basic and effective characteristics. The project applies VGG16 as its functional extractor, which enables the derivation of essential features from SNAP pictures of dangerous

software. The gathered tasks will serve as inputs to classify Malware samples into their family groups using the classification algorithm. The complex pattern combined with understanding conditions in VGG16S Pix makes it a key feature, which enhances version accuracy in the Malware classification process.

CNN

CNN functions as a deep learning framework to analyze structured data including photographs because of its engineered deep learning framework capabilities. The venture employs CNN[16] as an impartial classifier for the categorization of malware households. It has numerous layers, including convolutional, pooling, and completely linked layers, which facilitate the automatic learning of hierarchical features from enter records. Through the evaluation of malware pictures, the CNN acquires the ability to perceive pertinent functions and expect the malware own family related to every pattern. The ability to seize spatial correlations in pix renders CNN a formidable tool for unique malware classification.

Xception

Xception is a deep learning model architecture developed by Google studies, recognized for its efficacy and effectiveness in photograph classification duties. Xception is utilized in the project as a feature extractor to get vast capabilities from malware snap shots. Those attributes are subsequently employed for the category of malware families. The architecture of Xception [17] improves feature extraction with the aid of the implementation of depthwise separable convolutions, facilitating efficient records go with the flow and minimizing the parameter be counted. Xception's potential to figure complex patterns and relationships in pics renders it an critical element within the malware classification system, enhancing the accuracy of classification results.

Modified Attention CNN

The modified attention CNN is a convolutional neural network architecture augmented with interest mechanisms to concentrate on sizeable regions of input facts. This approach is employed for the category of malware households making use of non-disassembled binary files inside the project. by along with attention processes, the model can efficiently prioritize features derived from malware pictures and structural entropies. This allows advanced feature integration and augmented classification precision. The modified attention CNN layout enhances the resilience and efficacy of the malware type machine by selectively that specialize in salient features, therefore augmenting the model's capacity to accurately differentiate among various malware families.

4. EXPERIMENTAL RESULTS

Accuracy: The accuracy of the test concerns its ability to correctly distinguish between patient and healthy cases. In order to assess the accuracy of the test, one must calculate the ratio of real positives and real negatives in all evaluated cases. This can be mathematically expressed as:

"Accuracy = TP + TN TP + TN + FP + FN".

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision: The accuracy evaluates the proportion of precisely categorized cases among cases identified as fine. As a result, the formula for calculating accuracy is expressed:

"Precision = True positives/ (True positives + False positives) = TP/(TP + FP)"

$$\text{Precision} = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

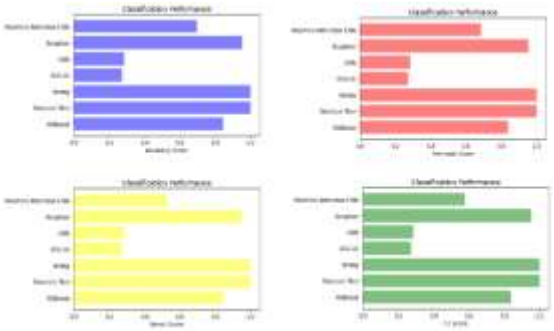
Recall: The reminder is a metric in "machine learning" that evaluates the ability of the model to recognize all the relevant times of a particular class. It is the ratio of precisely predicted positive observations to the whole real positives and offers knowledge directly to the effectiveness of the version when identifying the occurrence of a particular class.

$$Recall = \frac{TP}{TP + FN}$$

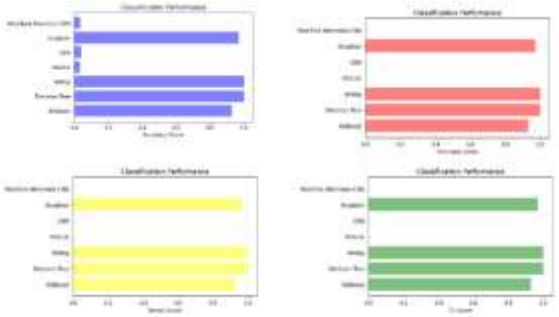
F1-Score: The F1 score is a metric for evaluating the accuracy of the machine-learning model. It integrates the metrics of accuracy and model. The metric of accuracy quantifies the frequency of real predictions generated by the model throughout the data file.

$$\textbf{F1 Score} = \frac{2}{\left(\frac{1}{\text{Precision}} + \frac{1}{\text{Recall}}\right)}$$

$$\textbf{F1 Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

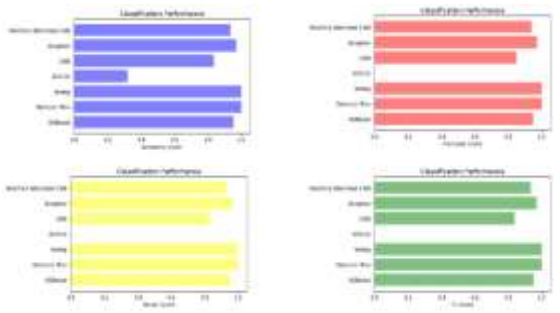


"Fig 3 COMPARISON GRAPHS OF BODMAS DATASET"



"Fig 4 COMPARISON GRAPHS OF BIG2015 DATASET"

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org



"Fig 5 COMPARISON GRAPHS OF MALIMG DATASET"

ML Model	Accuracy	Precision	Recall	F1_score
XGBoost	0.842	0.857	0.842	0.836
Extension Decision Tree	**1.000**	**1.000**	**1.000**	**1.000**
Extension Voting Classifier	**1.000**	**1.000**	**1.000**	**1.000**
VGG16	0.271	0.271	0.271	0.271
CNN	0.283	0.283	0.283	0.283
Extension Xception	**0.951**	**0.951**	**0.951**	**0.951**
Modified Attention CNN	0.694	0.680	0.525	0.577

"Fig 6 PERFORMANCE EVALUATION - BIG2015"

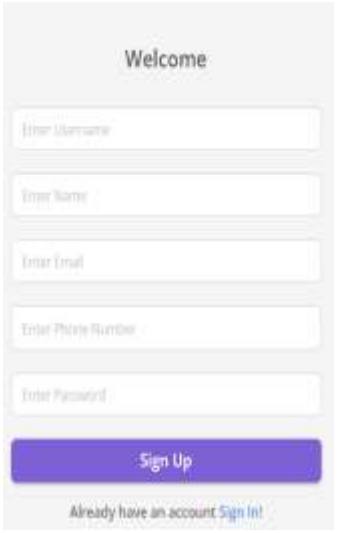
ML Model	Accuracy	Precision	Recall	F1_score
XGBoost	0.927	0.929	0.927	0.927
Extension Decision Tree	**1.000**	**1.000**	**1.000**	**1.000**
Extension Voting Classifier	**1.000**	**1.000**	**1.000**	**1.000**
VGG16	0.033	0.000	0.000	0.000
CNN	0.039	0.000	0.000	0.000
Extension Xception	**0.968**	**0.971**	**0.967**	**0.969**
Modified Attention CNN	0.037	0.000	0.000	0.000

"Fig 7 PERFORMANCE EVALUATION – BODMAS"

ML Model	Accuracy	Precision	Recall	F1_score
XGBoost	0.951	0.948	0.951	0.948
Extension Decision Tree	**1.000**	**1.000**	**1.000**	**1.000**
Extension Voting Classifier	**1.000**	**1.000**	**1.000**	**1.000**
VGG16	0.519	0.000	0.000	0.000
CNN	0.834	0.848	0.831	0.837
Extension Xception	**0.969**	**0.970**	**0.965**	**0.967**
Modified Attention CNN	0.954	0.936	0.931	0.933

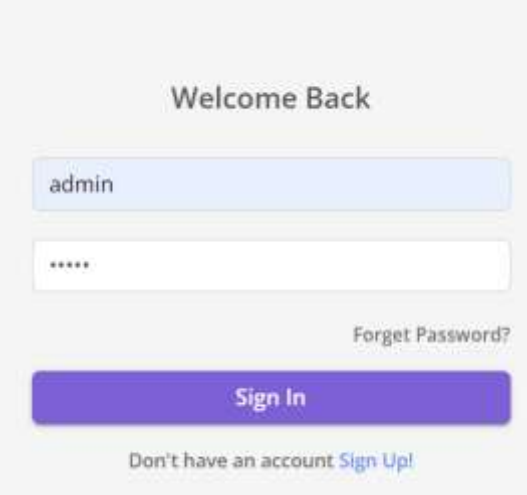
"Fig 8 PERFORMANCE EVALUATION – Malimg"

"Fig 9 Home Page"



"Fig 10 Sign Up"

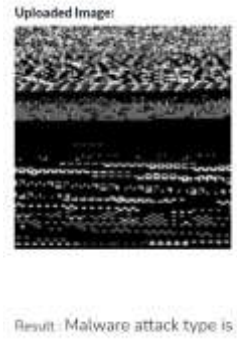
International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org



"Fig 11 Sign In"

"Fig 12 BIG2015"

"Fig 13 upload input images"

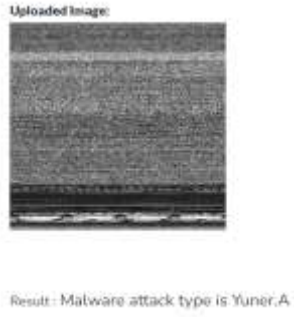


"Fig 14 predicted results"

"Fig 15 BODMAS"

"Fig 16 predicted results"

"Fig 17 MALIMG"



"Fig 18 predicted results"

Likewise, we can utilize alternative input data to forecast outcomes for the specified input data.

5. CONCLUSION

This study presents a CNN-based malware type model that significantly improves the accuracy of categorizing malware households with out requiring code disassembly. by using merging two modalities—malware pix and structural entropies—the model adeptly collects varying tiers of facts, hence improving classification overall performance. the integration of a cross-modal attention mechanism enhances and solidifies representations from both modalities, making certain coherent and thorough statistics representation. furthermore, the model's enhancement the usage of supplementary classifiers, including the "voting classifier" and decision Tree, attaining a hundred% accuracy, highlights its resilience and dependability. The incorporation of an intuitive Flask interface with sturdy authentication enhances both protection and value, rendering the system accessible and efficient for malware categorization activities. The recommended version and its expansions present a possible technique to the problems confronted via malware editions, ensuring a comfortable and user-pleasant environment for malware studies and categorization.

6. FUTURE SCOPE

The feature scope of the attention-based pass-Modal CNN utilizing Non-Disassembled files for Malware type includes numerous crucial factors. The version makes use of non-disassembled binary files as input facts, thereby obviating the onerous manner of code disassembly. This optimizes the classification technique and improves efficiency. Secondly, the use of a go-modal interest mechanism permits the version to successfully amalgamate input from two modalities: malware snap shots and structural entropies. This selection permits thorough characteristic integration, ensuring that both modalities considerably contribute to the class technique. The method seeks to reap specific category of malware families by using aligning and enhancing representations of malware images and structural entropies. By means of concurrently analyzing each modalities, the model can discern many attributes of malware samples, resulting in enhanced classification precision. The feature scope highlights the model's potential to utilize non-disassembled files and cross-modal attention for resilient and efficient malware categorization.

REFERENCES

[1] (2021). Picus Security. [Online]. Available: https://www.picussecurity. com/resource/blog/red-report-2021-top-ten-attack-techniques

[2] A. Shabtai, R. Moskovitch, Y. Elovici, and C. Glezer, ‘‘Detection of malicious code by applying machine learning classifiers on static features: A state-of-the-art survey,’’ Inf. Secur. Tech. Rep., vol. 14, no. 1, pp. 16–29, 2009. [Online]. Available:https://www.sciencedirect.com/science/article/pii/S1363412709000041

[3] A. Abusitta, M. Q. Li, and B. C. M. Fung, ‘‘Malware classification and composition analysis: A survey of recent developments,’’ J. Inf. Secur. Appl., vol. 59, Jun. 2021, Art. no. 102828. [Online].Available:https://www.sciencedirect.com/science/article/pii/S2214212621000648

[4] W. Han, J. Xue, Y. Wang, L. Huang, Z. Kong, and L. Mao, ‘‘MalDAE: Detecting and explaining malware based on correlation and fusion of static and dynamic characteristics,’’ Comput. Secur., vol. 83, pp. 208–233, Jun. 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S016740481831246X

[5] M. Sikorski and A. Honig, Practical Malware Analysis: The Hands-on Guide to Dissecting Malicious Software. San Francisco, CA, USA: No Starch Press, 2012.

[6] A. Afianian, S. Niksefat, B. Sadeghiyan, and D. Baptiste, ‘‘Malware dynamic analysis evasion techniques: A survey,’’ ACM Comput. Surveys, vol. 52, no. 6, pp. 1–28, Nov. 2020.

[7] M. Hassen, M. M. Carvalho, and P. K. Chan, ‘‘Malware classification using static analysis based features,’’ in Proc. IEEE Symp. Ser. Comput. Intell. (SSCI), Nov. 2017, pp. 1–7.

[8] H. Zhang, X. Xiao, F. Mercaldo, S. Ni, F. Martinelli, and A. K. Sangaiah, ‘‘Classification of ransomware families with machine learning based on N-gram of opcodes,’’ Future Gener. Comput. Syst., vol. 90, pp. 211–221, Jan. 2019. [Online]. Available:https://www.sciencedirect.com/science/article/pii/S0167739X18307325

[9] Hex Ray, IDA Pro-Hex Rays. Accessed: Mar. 7, 2023. [Online]. Available: https://www.hex-rays.com/ida-pro/

[10] D. Gibert, C. Mateu, and J. Planes, ‘‘HYDRA: A multimodal deep learning framework for malware classification,’’ Comput. Secur., vol. 95, Aug. 2020, Art. no. 101873. [Online]. Available:https://www.sciencedirect.com/science/article/pii/S0167404820301462

[11] D. Gibert, C. Mateu, and J. Planes, ‘‘Orthrus: A bimodal learning architecture for malware classification,’’ in Proc. Int. Joint Conf. Neural Netw. (IJCNN), Jul. 2020, pp. 1–8.

[12] X. Chong, Y. Gao, R. Zhang, J. Liu, X. Huang, and J. Zhao, ‘‘Classification of malware families based on efficient-net and 1D-CNN fusion,’’ Electronics, vol. 11, no. 19, p. 3064, Sep. 2022.

[13] D. Gibert, C. Mateu, J. Planes, and R. Vicens, ‘‘Using convolutional neural networks for classification of malware represented as images,’’ J. Comput. Virol. Hacking Techn., vol. 15, no. 1, pp. 15–28, Mar. 2019.

[14] M. Xiao, C. Guo, G. Shen, Y. Cui, and C. Jiang, ‘‘Image-based malware classification using section

distribution information,'' Comput. Secur., vol. 110, Nov. 2021, Art. no. 102420.

[15] D. Gibert, C. Mateu, J. Planes, and R. Vicens, ''Classification of malware by using structural entropy on convolutional neural networks,'' in Proc. AAAI Conf. Artif. Intell., 2018, pp. 1–6.

[16] S. Albawi, T. A. Mohammed, and S. Al-Zawi, ''Understanding of a convolutional neural network,'' in Proc. Int. Conf. Eng. Technol. (ICET), Aug. 2017, pp. 1–6.

[17] R. Ronen, M. Radu, C. Feuerstein, E. Yom-Tov, and M. Ahmadi, ''Microsoft malware classification challenge,'' 2018, arXiv:1802.10135.

[18] L. Nataraj, S. Karthikeyan, G. Jacob, and B. S. Manjunath, ''Malware images: Visualization and automatic classification,'' in Proc. 8th Int. Symp. Visualizat. Cyber Secur. New York, NY, USA: Association for Computing Machinery, Jul. 2011, pp. 1–7, doi: 10.1145/2016904.2016908.

[19] L. Yang, A. Ciptadi, I. Laziuk, A. Ahmadzadeh, and G. Wang, ''BODMAS: An open dataset for learning based temporal analysis of PE malware,'' in Proc. IEEE Secur. Privacy Workshops (SPW), May 2021, pp. 78–84.

[20] J. Kang, S. Jang, S. Li, Y.-S. Jeong, and Y. Sung, ''Long short-term memory-based malware classification method for information security,'' Comput. Elect. Eng., vol. 77, pp. 366–375, Jul. 2019.

[21] Y. Qiao, W. Zhang, X. Du, and M. Guizani, ''Malware classification based on multilayer perception and Word2Vec for IoT security,'' ACM Trans. Internet Technol., vol. 22, no. 1, pp. 1–22, Sep. 2021, doi: 10.1145/3436751.

[22] A. Bensaoud, N. Abudawaood, and J. Kalita, ''Classifying malware images with convolutional neural network models,'' Int. J. Netw. Secur., vol. 22, no. 6, pp. 1022–1031, Oct. 2020.

[23] D. Xue, J. Li, T. Lv, W. Wu, and J. Wang, ''Malware classification using probability scoring and machine learning,'' IEEE Access, vol. 7, pp. 91641–91656, 2019.

[24] R. Pascanu, J. W. Stokes, H. Sanossian, M. Marinescu, and A. Thomas, ''Malware classification with recurrent networks,'' in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), Apr. 2015, pp. 1916–1920.

[25] B. Athiwaratkun and J. W. Stokes, ''Malware classification with LSTM and GRU language models and a character-level CNN,'' in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), Mar. 2017, pp. 2482–2486.

[26] A. Pektas and T. Acarman, ''Malware classification based on API calls and behaviour analysis,'' IET Inf. Secur., vol. 12, no. 2, pp. 107–117, Mar. 2018.

[27] S. S. Hansen, T. M. T. Larsen, M. Stevanovic, and J. M. Pedersen, ''An approach for detection and family classification of malware based on behavioral analysis,'' in Proc. Int. Conf. Comput., Netw. Commun. (ICNC), Feb. 2016, pp. 1–5.

[28] D. Ramachandram and G. W. Taylor, ‘‘Deep multimodal learning: A survey on recent advances and trends,’’ IEEE Signal Process. Mag., vol. 34, no. 6, pp. 96–108, Nov. 2017.

[29] X. Xu, T. Wang, Y. Yang, L. Zuo, F. Shen, and H. T. Shen, ‘‘Cross-modal attention with semantic consistence for image-text matching,’’ IEEE Trans. Neural Netw. Learn. Syst., vol. 31, no. 12, pp. 5412–5425, Dec. 2020.

[30] I. J. Cruickshank and K. M. Carley, ‘‘Analysis of malware communities using multi-modal features,’’ IEEE Access, vol. 8, pp. 77435–77448, 2020.

[31] P. Velickovic, D. Wang, N. D. Lane, and P. Lio, ‘‘X-CNN: Cross-modal convolutional neural networks for sparse datasets,’’ in Proc. IEEE Symp. Ser. Comput. Intell. (SSCI), Dec. 2016, pp. 1–8.

[32] Y.-H. H. Tsai, S. Bai, P. P. Liang, J. Z. Kolter, L.-P. Morency, and R. Salakhutdinov, ‘‘Multimodal transformer for unaligned multimodal language sequences,’’ in Proc. 57th Annu. Meeting Assoc. Comput. Linguistics, Jul. 2019, pp. 6558–6569. [Online]. Available: https://aclanthology.org/P19-1656

[33] C. E. Shannon, ‘‘A mathematical theory of communication,’’ Bell Syst. Tech. J., vol. 27, no. 3, pp. 379–423, 1948. [Online]. Available: http://plan9.belllabs.com/cm/ms/what/shannonday/shannon1948.pdf

[34] J. Kim, E.-S. Cho, and J.-Y. Paik, ‘‘Poster: Feature engineering using file layout for malware detection,’’ in Proc. Annu. Comput. Secur. Appl. Conf., Dec. 2020.

[35] M.-T. Luong, H. Pham, and C. D. Manning, ‘‘Effective approaches to attention-based neural machine translation,’’ in Proc. EMNLP, Aug. 2015, pp. 1412–1421. [Online]. Available: https://aclanthology.org/D15-1166

[36] J. Yan, G. Yan, and D. Jin, ‘‘Classifying malware represented as control flow graphs using deep graph convolutional neural network,’’ in Proc. 49th Annu. IEEE/IFIP Int. Conf. Dependable Syst. Netw. (DSN), Jun. 2019, pp. 52–63.

[37] M. Mays, N. Drabinsky, and S. Brandle, ‘‘Feature selection for malware classification,’’ in Proc. MAICS, Apr. 2017, pp. 165–170.

[38] Y. Zhang, Q. Huang, X. Ma, Z. Yang, and J. Jiang, ‘‘Using multi-features and ensemble learning method for imbalanced malware classification,’’ in Proc. IEEE Trustcom/BigDataSE/ISPA, Aug. 2016, pp. 965–973.

[39] M. Ahmadi, D. Ulyanov, S. Semenov, M. Trofimov, and G. Giacinto, ‘‘Novel feature extraction, selection and fusion for effective malware family classification,’’ in Proc. 6th ACM Conf. Data Appl. Secur. Privacy, Mar. 2016, pp. 183–194.

[40] R. Mitsuhashi and T. Shinagawa, ‘‘Deriving optimal deep learning models for image-based malware classification,’’ in Proc. 37th ACM/SIGAPP Symp. Appl. Comput. New York, NY, USA: Association for Computing Machinery,

Apr. 2022, pp. 1727–1731, doi: 10.1145/3477314.3507242.

[41] J. H. Go, T. Jan, M. Mohanty, O. P. Patel, D. Puthal, and M. Prasad, ‘‘Visualization approach for malware classification with ResNeXt,’’ in Proc. IEEE Congr. Evol. Comput. (CEC), Jul. 2020, pp. 1–7.

[42] Y.-S. Liu, Y.-K. Lai, Z.-H. Wang, and H.-B. Yan, ‘‘A new learning approach to malware classification using discriminative feature extraction,’’ IEEE Access, vol. 7, pp. 13015–13023, 2019.

[43] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, and S. Venkatraman, ‘‘Robust intelligent malware detection using deep learning,’’ IEEE Access, vol. 7, pp. 46717–46738, 2019.