Multi-Modal Features Representation-Based Convolutional Neural Network Model for Malicious Website Detection

[1] Ms. Renuka B N [2] Priyanka N C

[1]Assistant professor, Department Of MCA, BIET, Davangere

[2]Student, 4th semester MCA, Department Of MCA, BIET, Davangere

ABSTRACT : Web applications have thrived across numerous business sectors, serving as essential tools for numerous of users in their daily lives activities. Though, several numbers of these applications are malicious which is major threat to Internet users as they can steal delicate information, connect malware, and propagate spam. Detecting malicious websites by analysing web content is ineffective due to the difficulty of extraction of the representative features, the huge data volume, the evolving nature of the malicious patterns, the stealthy nature of the attacks, and the limitations of traditional classifiers. Uniform Resource Locators (URL) features are static and can often provide immediate insights about the website without the need to load its content. However, prevailing solutions for sleuthing malicious web applications through web contented study often struggle due to complex feature extraction, massive data volumes, evolving attack patterns, and limitations of traditional classifiers. This study proposes a multimodal representation approach that fuses textual and image-based features to improve the recital of the malicious website detection. Textual features facilitate the deep learning model's ability to understand and represent detailed semantic information related to attack patterns, while image features are effective in recognizing more general malicious patterns. In doing so, patterns that are hidden in textual format may be recognizable in image format. Deuce Convolutional Neural Network (CNN) replicas were constructed to excerpt the veiled structures from together textual and image represented features. Outcomes illustration the efficiency of the proposed model when related to other models. The overall performance in relations of Matthews Correlation Coefficient (MCC) was improved by 4.3% although the false positive rate was reduced by 1.5%.

Keywords: Matthews Correlation Coefficient (MCC)

I. INTRODUCTION

The exponential growth of web applications across numerous sectors has brought significant benefits but also a surge in cybersecurity threats, particularly malicious websites. These websites can compromise user data, mount malware, and facilitate spam distribution. Traditional detection methods relying solely on etymological URL features or static content analysis have proven insufficient, primarily due to the dynamic and stealthy nature of modern cyber threats. The complex and evolving structure of malicious websites makes feature extraction and accurate classification particularly challenging. To discourse these restrictions, this learning proposes a new multimodal method that syndicates both textual and image-based feature representations. By leveraging the strengths of Convolutional Neural Networks (CNN) for feature extraction and Artificial Neural Network for last classification, the proposed system offers an effective and scalable solution for enhancing malicious website detection.

II. LITERATURE REVIEW

K. Rachmawati, M. Bukhori, F. Nuryanti, and S. Hidayatullah, "Collaboration technology acceptance model, subjective norms and personal innovations on buying interest online," *Int. J. Innov. Sci. Res. Technol.*,2020. The rapid advancement of information technology has permeated diverse sectors—including manufacturing, banking, healthcare, hospitality, agriculture—and has underscored the crucial role of the internet in meeting everyday human needs. From students to professionals, individuals increasingly rely on internet connectivity to access, share, and retrieve information swiftly [1].

C.I. Agustyaningrum, R. Aryanti, M. Haris, and T. Misriati, "Online shopper intention analysis using conventional ML and deep neural network classification algorithm," *Jurnal Penelitian Pos dan Informatika* Nov. 2021. The global proliferation of e-commerce in recent years has showcased its immense market potential, with rising sales signalling promising opportunities. To maximize store profitability, this study investigates the prediction and classification of online shopper intentions using both the conventional machine learning (ML) algorithms and deep neural networks (DNNs) [2].

J. A. Al-Gasawneh, M. H. Al-Wadi, B. M. Al-Wadi, B. E. Alown, and N. M. Nuseirat, "The interaction effect of comprehensiveness between social media and online purchasing intention in Jordanian pharmacies," *Int. J. Interact. Mobile Technol.*, Sep. 2020.

This experimental study examines how the comprehensiveness of social media content moderates the association between social media custom and online purchasing intentions among customers of Jordanian pharmacies. Drawing upon both the Technology Acceptance Model (TAM) and the Theory of Planned Behaviour (TPB), the authors conducted a survey with 198 respondents [3].

N. Provos, D. McNamee, P. Mavrommatis, K. Wang, and N. Modadugu, "The ghost in the browser: Analysis of web-based malware," *HotBots*, Apr. 2019. This seminal study examines the escalating threat posed by web-based malware—malicious software that infects users' computers simply through visiting compromised websites. Unlike classic push-based botnet attacks, these browser-based infections operate on a pull model [4].

K. Townsend, "18.5 Million websites infested with malware at all time," Wired Bus. Media, SecurityWeek, Boston, MA, USA, Tech. Rep. Q4 2017, 2022. Accessed: Feb. 1, 2022. In the fourth quarter of 2017, SiteLock's analysis of over six million websites revealed that approximately 1%—

or nearly 18.5 million sites globally—were diseased with malicious at any given time Despite a 20% year-over-year rise in infections (from 0.8% to just over 1%), the incidence of attack attempts actually fell by around 20%, suggesting that cybercriminals [5].

A. S. Raja, R. Vinodini, and A. Kavitha, "Lexical structures based spiteful URL detection using ML techniques," *Mater. Today*, Jan. 2021 .This study addresses the escalating threat of malicious URLs—web links crafted by cybercriminals to deceive users and unleash malware, steal credentials, or initiate unauthorized system access -by proposing a lightweight, machine learning-based detection approach that relies exclusively on lexical URL features [6].

A. Subasi, M. Balfaqih, Z. Balfagih, and K. Alfawwaz, "A qualified evaluation of collaborative classifiers for malicious webpage detection," *Proc. Comput. Sci.*, Jan. 2021 This study presents a comprehensive evaluation of ensemble ML techniques for sleuthing malicious webpages sites deliberately crafted or compromised to facilitate cybercrime. [7].

S. R. Zahra, M. A. Chishti, A. I. Baba, and F. Wu, "Detecting COVID-19 disorder driven phishing/malicious URL bouts by a fuzzy logic and data mining-based intelligence system," *Egyptian Informat. J.*, Jul. 2022 .This study explores how the COVID-19 pandemic—with its global uncertainty and abrupt shift to digital platforms—created ideal conditions for cybercriminals to launch widespread phishing, malware, ransomware, and identity-theft campaigns [8].

III. EXISTING SYSTEM

There are leash main methods that have been suggested by researchers for malicious URL classification: blacklist, content-based, and URL-based Many techniques were proposed to hypothesis the detection classifiers such as usage of heuristic rules based on professional experience or the use of ML techniques. However, effective malicious URL detection is still an open issue problem. The performance of the recent malicious website detection solutions is influenced by the mined topographies and the ML algorithms used for constructing the detection classifier. Authors in presented an in-depth literature review that covers various machine learning-based techniques for detecting malicious URLs, considering aspects such as limitations, detection technologies, feature types, and datasets. The type of extracted topographies joint with deep learning techniques are research trends of malicious website detection solutions. The professional experience heuristic rule was broadly used for making a blacklist of malicious URLs such as Google safe web browsing tool However, the blacklist solutions are ineffective for spiteful URL recognition due to the constantly evolving threats causing the essential for frequent identification of the evolved threat and frequently updating the database. In order to identify harmful content on websites, numerous researchers have employed feature extraction techniques. For representation, natural language processing has been widely used. However, because the methods used by attackers are constantly changing, malicious website content is complex and such patterns become dynamic and stealthy leading to poor detection accuracy. For example, in the authors investigated how malicious websites employ various web spam techniques to evade detection. The intention is to deliver an effective solution for detecting and combating malicious websites that utilize techniques like redirection spam, hidden I frames spam, and content hiding spam. Accordingly, the study focuses on capturing screenshots of webpages from a user's perspective and using Convolutional Neural Network for classification. However, the solution is limited for detecting spam techniques. Moreover, the feature depends on screenshots of the loaded page might be dangerous and uncompleted due to the dynamic nature of the websites. In the authors collected features from the HTTP/s responses and applied various feature transformation and selection techniques for classification. However, these features are dynamic, subject to obfuscation using encoding and encryption mechanisms, which can render the detection classifier ineffective. While ML algorithms were widely used for constructing the detection classifier, many researchers focused on deep learning techniques. Deep learning can accurately determine the similar patterns learned during the training resulting in effective classification. This is because malicious domains are generated algorithmically while benign domains are created by humans. Thus, malicious URLs may contain more prominent features compared to the topographies extracted from the content which container be obfuscated, or encrypted to mislead the learning process. Authors in focused on detecting the malevolent URLs that are engendered algorithmically.

DISADVANTAGES

In an Existing system, solutions for detecting malicious web applications through web content analysis often struggle due to complex feature extraction, massive data volumes, evolving attack patterns, and limitations of traditional classifiers. Relying solely on lexical URL features proves insufficient, potentially leading to inaccurate classifications.

IV. PROPOSED SYSTEM

The system recommends a innovative multimodal representation approach that integrates textual and image-based features to enhance malicious website detection. This approach leverages the strengths of both modalities: textual features capture detailed semantic information related to attack patterns, and image features recognize broader malicious visual cues. Hidden patterns within textual content may become discernible through image analysis.

The proposed method services two Convolutional Neural Networks (CNNs): one for textual features and another for image features. For better decision-making, their outputs are then combined and fed into a classifier that uses artificial neural networks. Our findings show that the suggested model outperforms current methods. We achieve a 4.3% increase in Matthews Correlation Coefficient (MCC) and a 1.5% reduction in the false-positive rate, showcasing the effectiveness of our multimodal approach in accurately identifying malicious web applications.

ADVANTAGES

1. Integrating DNS-derived features through URL based structures enhances the comprehensiveness of spiteful website detection. This synergy offers valuable contextual information regarding domain behaviour and infrastructure, thereby fortifying the evaluation of website authenticity and security contributing to a more robust and nuanced approach to identifying malicious websites.

2. The study introduces a multimodal representation approach that utilizes both textual and image-based features to represent a comprehensive feature set. Textual features facilitate the deep learning model's ability to understand and represent detailed semantic information related to attack patterns, while image features are effective in recognizing more general malicious patterns.

3. To extract hidden features from the textual and image representations, design and develop two Convolutional Neural Network (CNN) models.

4. An additional, deep learning classifier was constructed to learn the relationships among the hidden features extracted by the CNN replicas. By using deep learning techniques to integrate and utilize both textual and visual information for more efficient malicious website detection, this method advances the field.

System Architecture

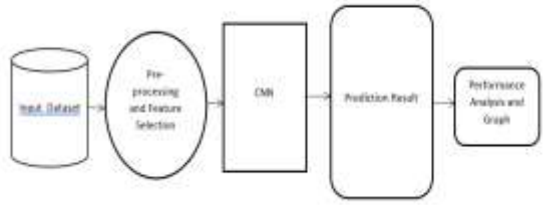
Input Dataset
Pre-processing and Feature Selection
CNN
Prediction Result
Performance Analysis and Graph

Fig 4.1. System Architecture

V. MODULE DESCRIPTION

The projected system is instigated as a web-based application using Python, Django, and deep learning libraries. It includes several interconnected modules that support the training, testing, and visualization of malicious website detection using a multimodal CNN model. Below are the key components and functionalities:

1. Server Login Module

- Provides authenticated access to system features for administrators and researchers.
- Ensures secure usage of dataset management, training, and prediction functionalities.
- Enables user role management and session control for accessing restricted modules.

2. Dataset Browser Module

- Allows users to browse and upload datasets for training and testing.
- Supports various data formats and previews the structure of textual and image-based data.
- Offers dataset splitting for model training & validation purposes (e.g., 80/20 or 70/30).

3. Train and Test Module

- Trains the multimodal CNN model using combined textual and image features from malicious and benign website datasets.
- There are two distinct CNN models in used
  - **Text-CNN**: For analyzing semantic patterns in URLs and webpage text.
  - **Image-CNN**: For analyzing screenshots or visual representations of webpages.
- The feature trajectories from mutually CNNs are fused and passed into Artificial Neural Network (ANN) for classification.
- Testing evaluates the model performance against unseen data to ensure generalization.

4. Accuracy Visualization (Bar Chart)

- Displays the training & testing accuracy of the prototypical using interactive bar charts.
- Helps in comparing the performance of different models or dataset versions.
- Shows metrics like accuracy, precision, F1-score, recall, and Matthews Correlation Coefficient (MCC).

5. Trained and Tested Accuracy Results

- Presents detailed numeric results after training and testing, including:
  - Confusion matrix
  - False Positive Rate (FPR)
  - True Positive Rate (TPR)
  - Matthews Correlation Coefficient (MCC)
- Helps evaluate model robustness and effectiveness in real-world scenarios.

6. Tweet-Type Prediction Module

- Enables testing the model on sample tweet data or real-time social media feeds for cybersecurity research.
- Predicts the type of content (e.g., suspicious, phishing, spam, benign).
- Demonstrates the model's flexibility and application to short-text environments like tweets.

7. Tweet-Type Graph Visualization

- Provides a pictorial illustration of tweet-type predictions.
- Graphs categorize tweets founded on their predicted label (e.g., benign vs. suspicious).
- Useful for showcasing real-time threat detection on social media platforms.

VI. RESULT

The recital of the projected multimodal detection model was evaluated against existing traditional and deep learning approaches. The model demonstrated significant improvements in detection accuracy, showcasing a 4.3% increase in the Matthews Correlation Coefficient (MCC), a key routine measured for binary cataloguing tasks. Furthermore, the false positive rate was condensed by 1.5%, indicating a more reliable classification process with fewer benign sites being misclassified as malicious. The integration of both image and textual features contributed to a more comprehensive understanding of the hidden malicious patterns, improving the model's generalizability and robustness. These results validate the effectiveness of using CNN-based multimodal feature fusion in identifying and mitigating malicious website threats.

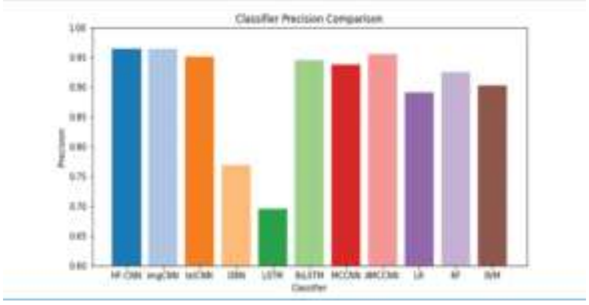

Fig 6.1. classifier precision comparison

VII. CONCLUSION

In conclusion, the proposed multimodal CNN-based detection model significantly enhances the capability of identifying malicious websites by leveraging both written and image-based features. This approach overcomes the limitations of traditional single-feature models and addresses the challenges posed by evolving cyberattack patterns. The fusion of semantic-rich textual data with visually observable image patterns allows for profounder vision into malicious behaviour. The improved Matthews Correlation Coefficient and lower false positive rate affirm the model's superior performance and practical applicability. This study not only shows that deep learning can be used to extract multimodal features, but it also offers a basis for future advancements in intelligent and adaptive cybersecurity systems.

VIII. REFERENCES

1. K. Rachmawati, M. Bukhori, F. Nuryanti, and S. Hidayatullah, "Collaboration technology acceptance model, subjective norms and personal innovations on buying interest online," *Int. J. Innov. Sci. Res. Technol.*, vol. 5, no. 11, pp. 115–122, 2020.

2. C. I. Agustyaningrum, R. Aryanti, M. Haris, and T. Misriati, "Online shopper intention analysis using conventional machine learning and deep neural network classification algorithm," *Jurnal Penelitian Pos dan Informatika*, vol. 11, no. 1, pp. 89–100, Nov. 2021.

3. J. A. Al-Gasawneh, M. H. Al-Wadi, B. M. Al-Wadi, B. E. Alown, and N. M. Nuseirat, "The interaction effect of comprehensiveness between social media and online purchasing intention in Jordanian pharmacies," *Int. J. Interact. Mobile Technol.*, vol. 14, no. 15, p. 208, Sep. 2020.

4. N. Provos, D. McNamee, P. Mavrommatis, K. Wang, and N. Modadugu, "The ghost in the browser: Analysis of web-based malware," *HotBots*, vol. 7, p. 4, Apr. 2019.

5. K. Townsend, "18.5Million websites infested with malicious at any time," Wired Bus. Media, SecurityWeek, Boston, MA, USA, Tech. Rep. Q4 2017, 2022. Accessed: Feb. 1, 2022.

6. A. S. Raja, R. Vinodini, and A. Kavitha, "Lexical features based malicious URL detection using machine learning techniques," *Mater. Today*, vol. 47, pp. 163–166, Jan. 2021

7. A. Subasi, M. Balfaqih, Z. Balfagih, and K. Alfawwaz, "A relative evaluation of ensemble classifiers for malevolent webpage detection," *Proc. Comput. Sci.*, vol. 194, pp. 272–279, Jan. 2021, doi: 10.1016/j.procs.2021.

8. S. R. Zahra, M. A. Chishti, A. I. Baba, and F. Wu, "Detecting COVID-19 bedlam driven phishing/malicious URL bouts by a fuzzy logic and data mining-based intelligence system," *Egyptian Informat. J.*, vol. 23, no. 2, pp. 197–214, Jul. 2022