Malware Detection in JPEG

Prof. F. S. Ghodichor[1], **Prashant Kharche**[2], **Chaitanya Katore**[3], **Ajit Adavale**[4], **Dipkumar Prajapat**[5]
Assistant Professor, Department of Information Technology[1]
Students, Department of Information Technology[2,3,4,5]
Sinhgad Institute of Technology, Lonavala, Maharashtra, India

Abstract*:* *Cyberattacks on people, companies, and organizations have grown in frequency. Cybercriminals are constantly searching for efficient ways to infect targets with malware in order to initiate an attack. Millions of people use images every day all throughout the world, and the majority of users think pictures to be secure for usage, however some kinds of pictures have the potential to carry a malware payload and execute detrimental acts. The main reason JPEG is the most widely used image format is because of its lossy compression. It's applied almost everyone, from small businesses to major corporations, and is present on nearly all devices (on digital cameras, cellphones, social networking, websites, etc.). Due of their reputation for being innocuous, enormous JPEG images have a lot of potential for misuse.*

Keywords: JPEG, Automatic Interpretation, Image Processing, Artificial Intelligence, Malware Detection, CNN, Deep learning

I. INTRODUCTION

Malware, or malicious software, is software created to infect a machine without the user's knowledge or consent. It is actually a generic definition for all sorts of threats that can affect a computer. A simple bracket of malware consists of train infectors and stand- alone malware. The objectives of a malware could include accessing private networks, stealing sensitive data, taking over computer systems to make use of its resources, or disrupting computing or communication operations.

1.1 MOTIVATION

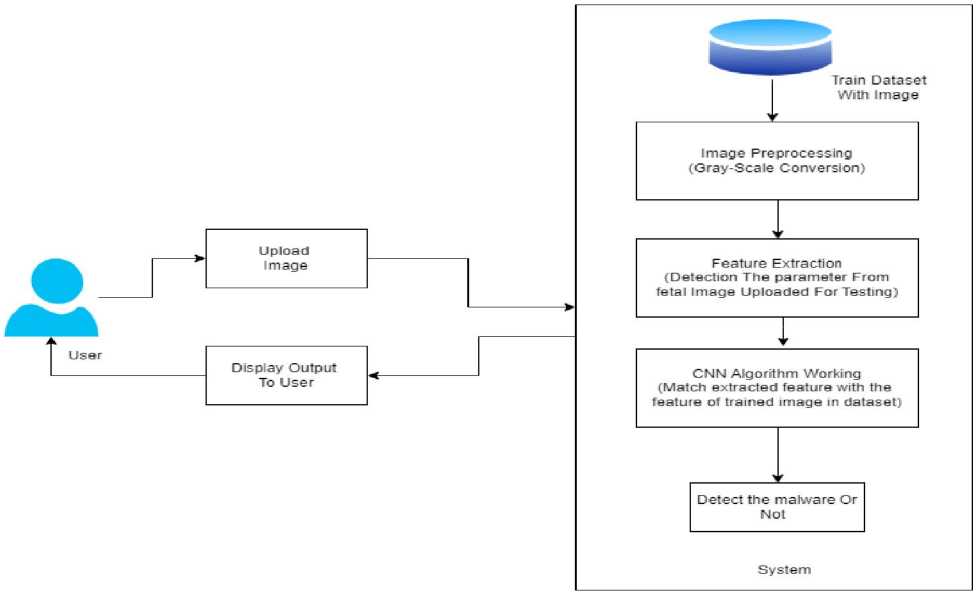
Embarking on the Malware Detection in JPEG files project presents a unique and impactful opportunity to contribute to the ever-evolving landscape of cybersecurity. By addressing a specific and often overlooked vector of attack, you have the chance to make a significant difference in enhancing digital security. Motivation for this project lies in the potential to develop specialized algorithms that can discern concealed threats within JPEG files, thereby fortifying existing cybersecurity measures. As you delve into the intricacies of image-based malware detection, envision the real-world applications – safeguarding digital image repositories, protecting e-commerce platforms, and securing critical infrastructure.

1.2 OBJECTIVE

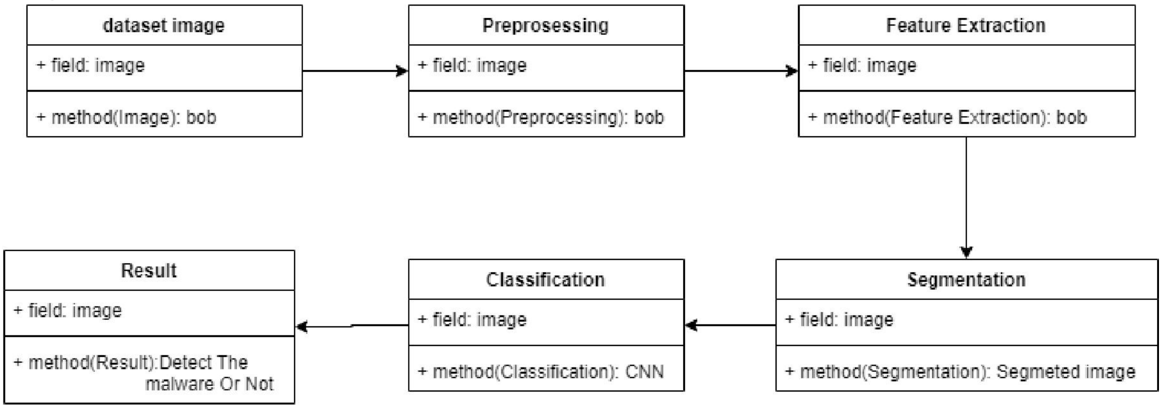
The objectives of the Malware Detection in JPEG files project are multi-faceted, aiming to develop specialized detection algorithms that distinguish between benign and malicious content within JPEG files. The primary goal is to enhance the accuracy and precision of malware detection by refining feature extraction methods and analysis techniques, thereby minimizing false positives and negatives. Addressing data limitations is crucial for training robust algorithms, necessitating the curation of diverse datasets

DOI: 10.48175/IJARSCT-13849

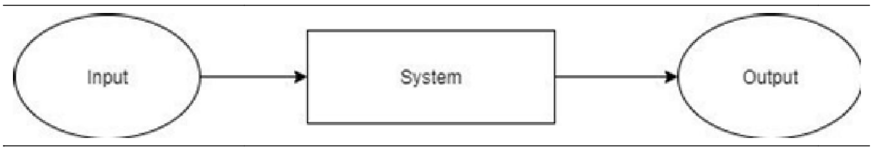
II. SYSTEM ARCHITECTURE



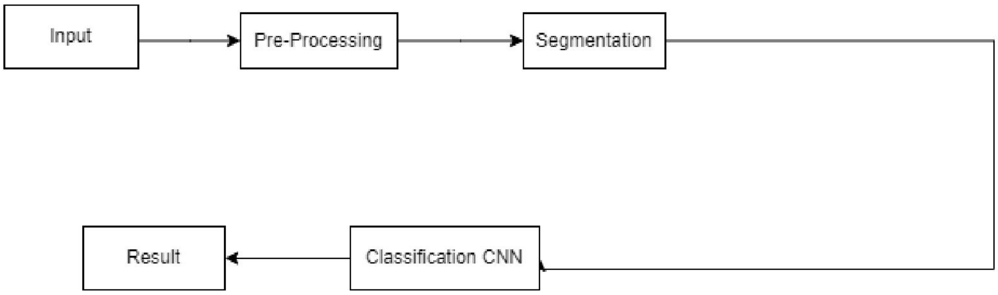
UML Diagram:



Data Flow Diagrams:

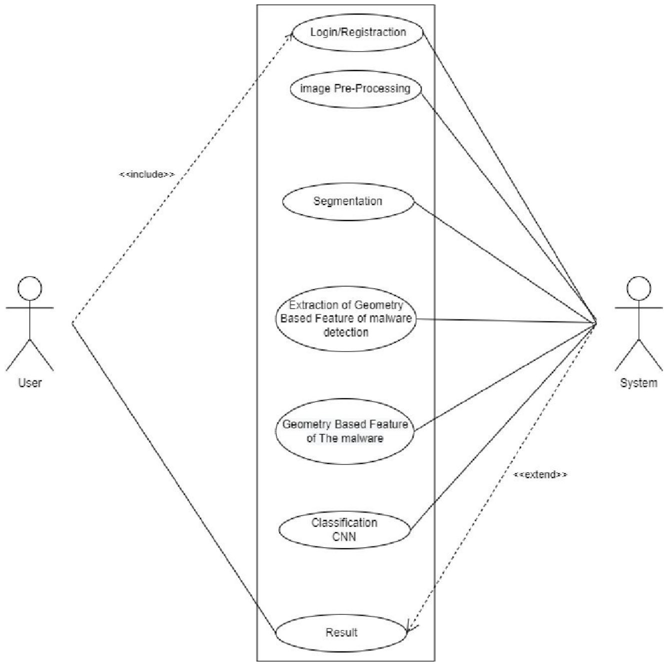


Data flow (0) diagram

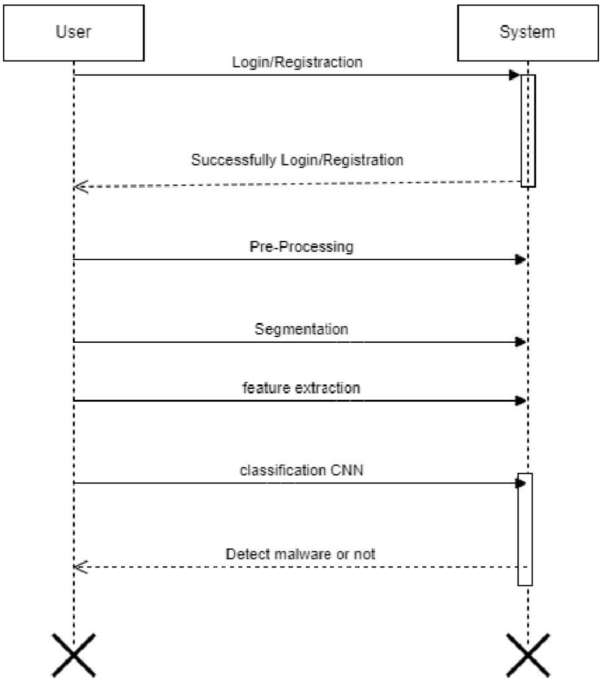


Dataflow (2) diagram

Use case and Sequence Diagrams:



Use case diagram



Sequence diagram

III. SOFTWAREREQUIREMENTS

- OperatingSystem:Windows10
- IDE:Spyder
- Programming Language: Python

IV. HARDWAREREQUIREMENTS

- Hardware: intelcore
- Speed:2.80GHz
- RAM:8GB
- HardDisk:500GB

V. APPLICATIONS

- Defense for E-commerce Platform.
- Protection for Digital Image Repositories.
- Securing Communication Channel.

VI. CONCLUSION

In conclusion, the Malware Detection in JPEG files project represents a significant stride towards fortifying digital security in the face of emerging threats. By addressing the specific challenge of malware concealed within JPEG files, the project introduces a specialized layer of defense that complements traditional cybersecurity measures. The enhanced accuracy, adaptability to evolving tactics, and contributions to threat intelligence make it a valuable asset for a wide range of applications, from safeguarding digital image repositories and e-commerce platforms to securing critical infrastructure.

VII. ACKNOWLEDGMENT

It gives great pleasure to present the preliminary project report on the project topic, "Malware Detection in JPEG." We take this opportunity to thank our internal guide, Prof. F. S. Ghodichor, for giving us all the help and guidance we needed. We're thankful to him for his kind support. His valuable suggestions were quite helpful.

REFERENCES

[1].pp. 7–11, 2014. 2. E. S. Solutions and Q. Heal, “Quick Heal Quarterly Threat Report | Q1 2017,” 2017 url:http://www.quickheal.co.in/resources/threat-reports . [Accessed: 13-june-2017].

[2]. A. Govindaraju, “Exhaustive Statistical Analysis for Detection of Metamorphic Malware,” Master’s project report, Department of Computer Science, San Jose State University, 2010.

[3]. M. G. Schultz, E. Eskin, and S. J. Stolfo, “Data Mining Methods for Detection of New Malicious Executables,” 2001.

[4]. D. Bilar, “Opcodes As Predictor for Malware,” International Journal of Electronic Security and Digital Forensics, vol. 1, no. 2, pp. 156–168, 2007.

[5]. Y. Elovici, A. Shabtai, R. Moskovitch, G. Tahan, and C. Glezer, “Applying Machine Learning Techniques for Detection of Malicious Code in Network Traffic,” Annual Conference on Artificial Intelligence. Springer Berlin Heidelberg, pp. 44–50, 2007.

[6]. R. Moskovitch, D. Stopel, C. Feher, N. Nissim, N. Japkowicz, and Y. Elovici, “Unknown malcode detection and the imbalance problem,” Journal in Computer Virology, vol. 5, no. 4, pp. 295–308, 2009.

[7].R. Moskovitch et al., “Unknown malcode detection using OPCODE representation,” Intelligence and Security Informatics. Springer Berlin Heidelberg, vol. 5376 LNCS, pp. 204–215, 2008

[8]. I. Santos, J. Nieves, and P. G. Bringas, “Semi-supervised learning for unknown malware detection,” International Symposium on Distributed Computing and Artificial Intelligence. Springer Berlin Heidelberg, vol. 91, pp. 415–422, 2011.

[9]. I. Santos, F. Brezo, X. Ugarte-Pedrero, and P. G. Bringas, “Opcode sequences as representation of executables for data-miningbased unknown malware detection,” Information Sciences, vol. 231, pp. 64–82, 2013.

[10]. A. Shabtai, R. Moskovitch, C. Feher, S. Dolev, and Y. Elovici, “Detecting unknown malicious code by applying classification techniques on OpCode patterns,” Security Informatics, vol. 1, no. 1, p. 1, 2012.

[11]. A. Sharma and S. K. Sahay, “An effective approach for classification of advanced malware with high accuracy,” International Journal of Security and its Applications, vol. 10, no. 4, pp. 249–266, 2016.

[12]. S. K. Sahay and A. Sharma, “Grouping the Executables to Detect Malwares with High Accuracy,” Procedia Computer Science, vol. 78, no. June, pp. 667–674, 2016.

[13].Kaggle, “Microsoft Malware Classification Challenge (BIG 2015)” Microsoft, URL: https://www.kaggle.com/c/malware-classification , [Accessed : 10/December/2016].

[14]. A. Sharma and S. K. Sahay, “Evolution and Detection of Polymorphic and Metamorphic Malware: A Survey,” International Journal of Computer Application, vol. 90, no. 2, pp. 7–11, 2014