Rafael Angelo Christianto

Made By: Rafael Angelo Christianto

Introduction

Sometimes attackers like to hide malware originally called “malware.exe”. But after publishing they usually disguise it as “cute_dogs_pics.jpg”. However, even if they already did that, the magic numbers of the file will still remain the same. The unique byte signatures stored at the beginning of every file type, cannot be altered without corrupting the file format.

Because of this, magic number analysis remains one of the most reliable methods for detecting file masquerading, extension spoofing, and suspicious file behavior.

This project aims to simplify the process by providing a web-based file extension scanner capable of identifying true file types, calculating cryptographic hashes, performing entropy analysis, and previewing the initial byte structure. Through these analyses, users can quickly determine whether a file matches its claimed format or has been tampered with.

Methodology

The approach focuses on analyzing uploaded files using a combination of lightweight yet essential tools and techniques. The core components include:

- **Magic Number Detection**
  The application reads the first few bytes of the file and compares them against known file signatures to determine the actual file type.=
- **Cryptographic Hashing**
  The scanner computes MD5, SHA-1, and SHA-256 hash values, providing integrity checks and compatibility with malware databases.
- **Entropy Calculation**
  Entropy is measured in bits per byte to identify anomalies such as encrypted or packed data.
- **Hex Preview Extraction**
  The first 256 bytes of the file are shown to assist manual inspection and verification of structural headers.

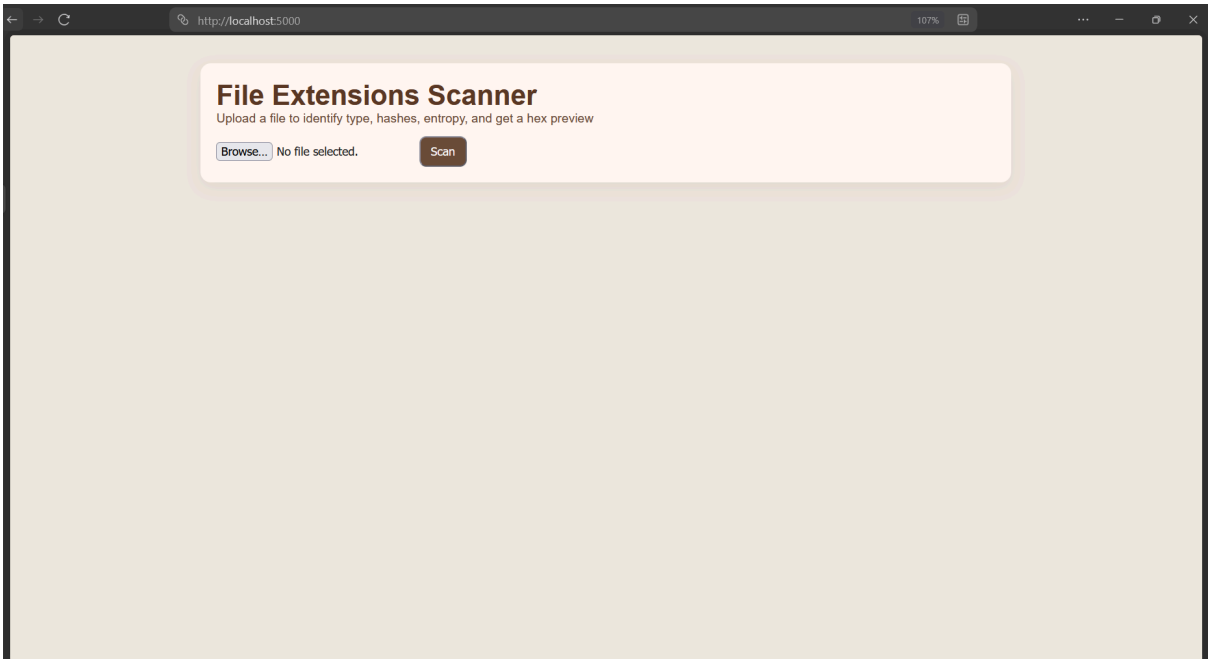
- **Flask Web Framework**
  The entire tool runs through a simple Flask-based web interface, allowing users to upload and analyze files directly from their browser.

Only essential libraries are used to maintain simplicity, reliability, and transparency in the analysis workflow.

Results

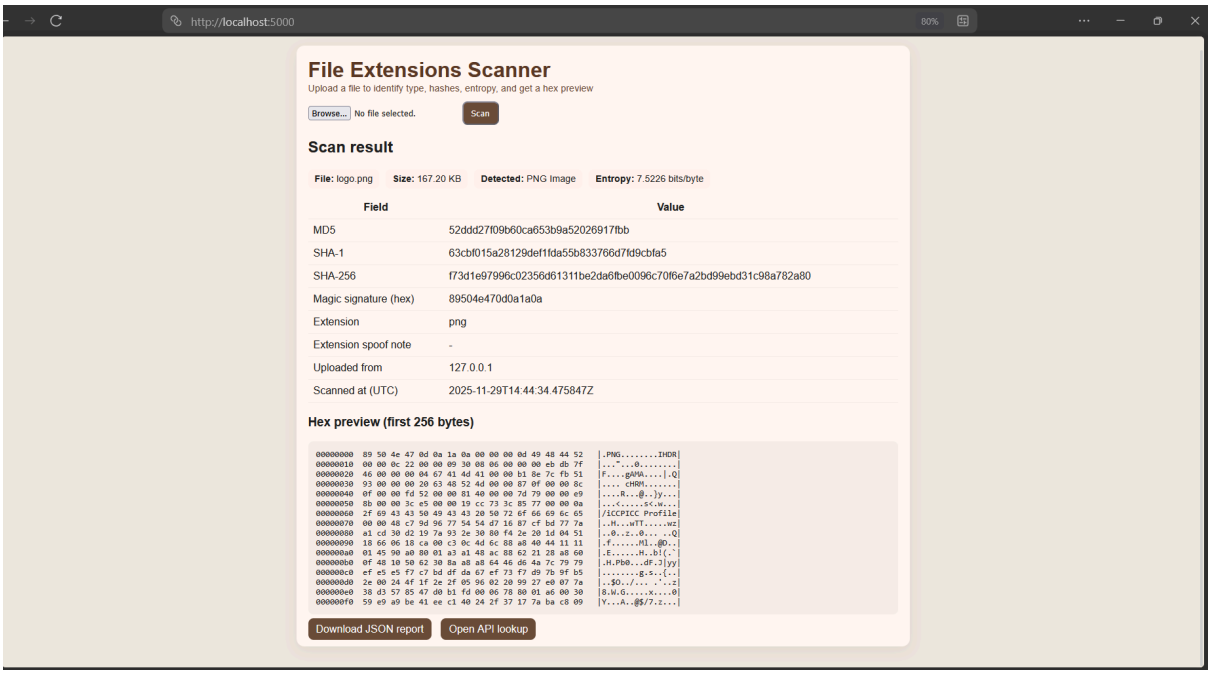
The code has been pushed to a GitHub repository at the following link: https://github.com/RafaelAngeloChristianto/file_extension_scanner

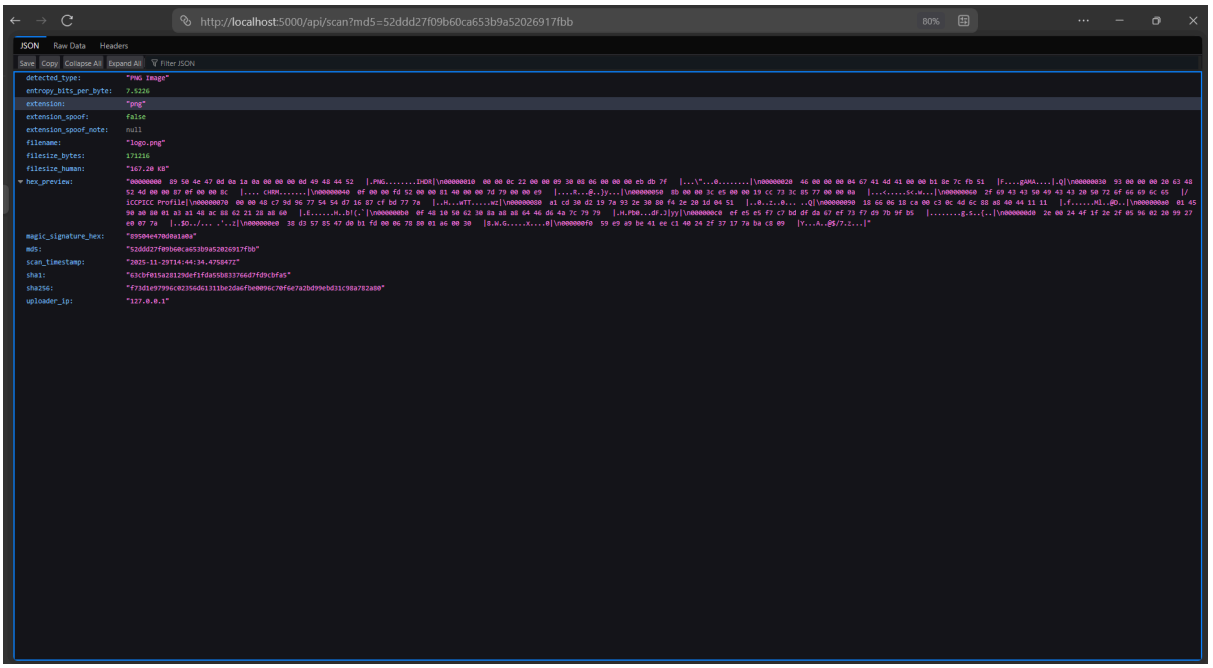
Upon running the program, the user has to go to the browser and type in localhost:5000, since that is the default local port for Flask applications. And they will be greeted with this simple GUI.



From the screenshots below we can see all the different hashes for the file, magic number, and the first 256 bytes of the hex preview. And for more options,

there is a JSON download option and Open API lookup feature, like shown in the second screenshot below.





Discussion

1. Overview

The file logo.png was analyzed to determine its type, structure, cryptographic properties, entropy, and initial byte sequence. All results indicate that the file is

a standard, non-malicious PNG image with no signs of tampering or extension spoofing.

2. File Metadata

Filename: logo.png
File size: 167.20 KB
Detected type: PNG Image
Upload source: 127.0.0.1
Scan time (UTC): 2025-11-29T14:51:23.717234Z
The detected file type matches the filename extension, and no inconsistencies were observed.

3. Cryptographic Hashes

The following hashes serve as unique fingerprints of the file:

- MD5: 52ddd27f09b60ca653b9a52026917fbb
- SHA-1: 63cbf015a28129def1fda55b833766d7fd9cbfa5
- SHA-256: f73d1e97996c02356d61311be2da6fbe0096c70f6e7a2bd99ebd31c98a782a80

These values can be used for integrity verification, duplication checks, and comparison against known malware signatures.

4. Magic Number / File Signature

Hex signature: 89504e470d0a1a0a
This corresponds to the official PNG signature defined in ISO/IEC 15948. The presence of this exact sequence confirms that the file is structurally a PNG and not an altered or disguised file.

5. Entropy Analysis

Entropy: 7.5226 bits/byte
This entropy level is normal for PNG images due to their compressed and visually complex data. No indication of abnormal compression, embedded payloads, or encrypted content was found based solely on entropy.

6. Extension Verification

Extension: png
Extension spoof note: None
The extension and magic number match. There is no sign of extension spoofing or inconsistent metadata.

7. Hex Preview Analysis (First 256 Bytes)

A preview of the first 256 bytes confirms the presence of standard PNG chunks:

- PNG signature
- IHDR (Image Header)
- gAMA (Gamma)
- cHRM (Chromaticity)
- iCCP (ICC Color Profile)

The data structure follows the correct PNG format, with no abnormalities such as appended executable code or non-standard ancillary chunks.
The IHDR content indicates:

- Image width encoded as: 0x00000C22 (3106 pixels)
- Image height encoded as: 0x00000930 (2352 pixels)

All observed bytes conform to expected PNG layout.

8. File Conclusion

No signs of malicious modification, embedded content, or structural anomalies were detected. The file appears to be a standard PNG image containing typical metadata and color profile information.
The file is considered safe based on:

- Matching signature and extension
- Normal entropy level
- Valid PNG chunk structure
- No suspicious appended data in the preview
- No hash matches indicating known malicious content (based on provided data only)