

Mini Project Report

on

“ Design and develop a tool for digital forensic of images”

Submitted by

Mr. Pratik Lonare **14232**

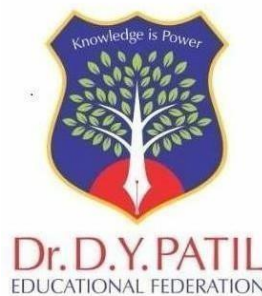
Mr. Kadayya Mathapati 14234

Mr. Rushikesh Patil **14244**

Mr.Anmol Pawar **14245**

of A.Y 2023-24

1stSemester, BE



Guided By

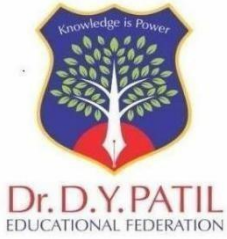
Prof. Sagar Dhanake

Department of Computer Engineering,

Dr. D. Y. Patil College of Engineering and Innovation

Varale, Talegaon, Pune

Affiliated to the Savitribai Phule Pune University



**DR. D.Y. PATIL COLLEGE OF ENGINEERING
AND INNOVATION**

**Affiliated to the Savitribai Phule Pune
University**



Certificate

This is to certify that **Mr. Pratik Lonare, Mr. Kadayya Mathapati, Mr. Rushikesh Patil, Mr. Anmol Pawar** Student of Final-year Computer Engineering of **DR. D.Y. Patil College of Engineering and Innovation** has successfully completed the Mini Project on “**Using online application analyze the image**” under the course **Laboratory Practice IV(Cyber Security And Digital Forensics) Mini Project (410242)** as prescribed in the curriculum for the academic year 2023 -2024.

Place: Talegaon, Pune

Prof. Sagar Dhanake
(Project Guide)

Dr. Alpana Adsul
(Head Of Department)

Dr.Suresh Mali
(Principal)

Seal Of Institute

Acknowledgement

It is great pleasure for me to acknowledge the assistance and contribution of number of individuals who helped me in Laboratory Practice III (Machine Learning) Mini Project (410242)

First and foremost, I wish to record my gratitude and thanks to Prof. Vishal Borate for his enthusiastic guidance and help in successful completion of Mini Project. I express my thanks to Dr. Alpana Adsul (Head of Computer Department), Dr. Suresh Mali (Principal) for their valuable guidance. I am also thankful to other teachers and non-teaching staff of Computer Engineering Department and Library for their co- operation and help.

THANKING TO ALL OF YOU...

**Mr. Pratik Lonare
Mr.Kadayya Mathapati
Mr. Rushikesh Patil
Mr. Anmol Pawar**

Index

Sr. No.	Content	Page No.
1	Abstract	5
2	Introduction	5
3	Problem Statement	6
4	Motivation	6
5	Objective	7
6	Implementation	8
7	Results	9
8	Conclusion	15
9	Reference	15

Abstract

In today's digitally-driven world, the analysis of images plays a pivotal role with broad-ranging applications spanning from content categorization to object recognition. This mini-project embarks on the development of an online Python application designed to harness image analysis techniques for processing and interpreting images sourced from the internet. The project's central focus lies in constructing an accessible and user-friendly platform where users can effortlessly upload images or input image URLs. The system, in turn, offers a spectrum of analyses, insights, and visualizations based on the content of these images. By creating this online Python application for image analysis, the mini-project seeks to empower users with a deeper understanding of image content, thus opening doors to a myriad of potential applications. Whether it's assisting photographers in the organization of their image collections or providing businesses with indispensable tools for content moderation and image recognition, this project underscores the immense power of Python in the realm of image analysis. Ultimately, it endeavors to present a valuable resource for those intrigued by the possibilities of image processing and machine learning techniques and showcases the remarkable potential Python holds in this field..

Introduction

In the digital era, the vast volume of images available on the internet has become an invaluable source of information and a playground for innovation. Images, whether they are shared on social media, published on websites, or stored in vast databases, contain a wealth of visual content that can be harnessed for various applications, from content classification to object recognition. As the saying goes, "A picture is worth a thousand words." Indeed, images often convey information that is difficult to articulate in text form. Whether it's understanding the composition of a photograph, detecting objects within an image, or classifying images into predefined categories, the ability to analyze images is a transformative capability with countless applications across diverse fields.

This mini project embarks on a journey to harness the power of Python in image analysis by creating an online application for analyzing images from the web. This project aims to make this power accessible to a wider audience by developing an online application that simplifies the process of image analysis. Through this application, users will be able to upload their own images or input image URLs from the internet, and in return, they will receive insightful analysis and visualizations of the image content. Leveraging the rich ecosystem of Python libraries, machine learning models, and web frameworks, this application is designed to bridge the gap between advanced image analysis techniques and non-technical users.

Problem Statement

In the realm of digital forensics, the burgeoning volume of images has presented a formidable challenge. The proliferation of visual content on the internet and across digital devices has given rise to the urgent need for an advanced tool for the comprehensive forensic analysis of images. Traditional methods often prove insufficient in efficiently extracting vital information and insights from this vast visual data. Consequently, there exists a significant gap in the tools available to digital forensics experts, cybersecurity professionals, and law enforcement agencies for tackling image-related evidence in an era where images play a pivotal role in investigations. This problem statement underscores the demand for the design and development of a robust tool capable of addressing the intricate challenges posed by image forensics, thereby enhancing the capabilities of digital forensics in today's data-intensive environment.

Motivation

The impetus behind embarking on the development of an online application for image analysis in Python stems from recognizing the ever-increasing relevance of visual data in the realm of cyber security and digital forensics. Images have become a ubiquitous form of communication, and they hold a significant place in our personal and professional lives. The following key factors drive the pursuit of this mini project within the context of cyber security and digital forensics:

- a) **Proliferation of Visual Data in Cyber Threats:** In the digital landscape, the proliferation of visual content is not only evident in benign use but also in malicious activities. Cybercriminals increasingly employ images to conceal malicious code, exfiltrate data, or deliver phishing attacks. Understanding, analyzing, and effectively countering these visual elements is crucial in safeguarding digital environments.
- b) **Evidence Extraction and Analysis:** Digital forensics often involves the examination of visual content as evidence in legal proceedings. Images can serve as essential pieces of the puzzle when investigating cybercrimes, and robust image analysis tools are essential for extracting pertinent information and drawing insights to support legal cases.
- c) **Accessibility for Digital Forensics Professionals:** Traditional image analysis tools are often complex and require significant expertise. Making image analysis accessible to digital forensics experts, who may not have in-depth computer vision skills, is paramount. This project aims to provide an accessible solution tailored to their needs.
- d) **Leveraging Technological Advances:** Advancements in machine learning, deep learning, and computer vision have revolutionized the way we can analyze images in a digital forensics context. These technological breakthroughs have opened up new avenues for efficiently scrutinizing visual data and identifying patterns and anomalies in digital investigations.
- e) **Multifaceted Applications in Cyber Security:** Image analysis plays a crucial role in various aspects of cyber security, from detecting malware within images to identifying suspicious patterns in network traffic. This project is envisioned as a versatile tool that can aid cyber security professionals in a range of tasks.

f) **Democratizing Digital Forensics:** This project aims to democratize digital forensics tools and techniques, making them accessible to a wider audience, including law enforcement agencies, cybersecurity teams, and digital forensics experts. It seeks to reduce barriers to entry and enable a broader set of professionals to conduct effective image analysis in digital investigations.

g) **Educational Value for Cyber Security Professionals:** In addition to its practical applications, this mini project serves as an educational platform for those in the field of cyber security and digital forensics. It offers an opportunity for professionals and aspiring digital forensics experts to gain hands-on experience in applying image analysis techniques to real-world cases, facilitating skills development and learning in this critical domain.

Objective

1. **Digital Evidence Extraction:** Extract a wide-ranging array of digital evidence from images relevant to cyber security incidents and digital forensic investigations. This includes identifying hidden data, metadata, timestamps, and any potentially malicious content concealed within the images.
2. **Forensic Image Analysis Models:** Implement cutting-edge machine learning models and algorithms tailored to the specific needs of cyber security and digital forensics. These models should be adept at tasks such as steganalysis, malware detection in images, forgery detection, and other forensic analyses.
3. **Real-Time Digital Investigation Support:** Facilitate real-time analysis of digital evidence, enabling investigators to swiftly ascertain the nature of threats, identify potential vulnerabilities, and gather evidence in an efficient and responsive manner. Real-time analysis is crucial for rapid incident response in cyber security and forensic investigations.
4. **Visual Representation of Digital Artifacts:** Provide clear, concise, and visually informative representations of digital artifacts and findings. This includes timeline visualizations of events, visual logs of activities, and graphical summaries of key findings that assist cyber security professionals and digital forensics experts in understanding the digital landscape and the context of their investigations.

Implementation

Extract text from image :

```
from PIL import
Imageimport

pytesseract

# Load the image
image = Image.open('C:\\Users\\rushi\\OneDrive\\Desktop\\gory_monalisas_feat_free.jpg')

try:
    extracted_text = pytesseract.image_to_string(image, config='--psm 6')except
Exception as e:
    print(f"An error occurred: {e}")

# Save to a text file
with open('extracted_text.txt', 'w', encoding='utf-8') as file:file.write(extracted_text)
```


Output:

≡ extracted_text.txt

```
1   Fee ere caummael Seer
2   Ld ye
3   ae ro uy - 2
4   AN ese - 2 > 4 % eels
5   "ee ae, _ _ ae
6   aes 73 1) ee , ge
7   (aes ip 4 ; .
8
9   a! (" \ Yast
10
11  ISS =
12
13  om
14
```

Metadata of Image :

```
from PIL import Image

from PIL.ExifTags import TAGS, GPSTAGS

def extract_metadata(image_path):
    try:
        image = Image.open(image_path)
        exif_data = image._getexif()

        if exif_data is not None:
            for tag, value in exif_data.items():
                tag_name = TAGS.get(tag, tag)
                print(f"{tag_name}: {value}")
    except Exception as e:
        print(f"An error occurred: {e}")

if __name__ == "__main__":
    image_path = 'C:\\Users\\rush\\OneDrive\\Desktop\\gory_monalisas_feat_free.jpg' #
    Replace with your image file path
    extract_metadata(image_path)
```

Output :


```
PS C:\Users\rushi\OneDrive\Desktop\python> python -u "c:\Users\rushi\OneDrive\Desktop\pytho
n\CS2.py"
ResolutionUnit: 2
ExifOffset: 164
Software: Adobe Photoshop CS5.1 Windows
Orientation: 1
DateTime: 2014:04:29 19:14:53
XResolution: 300.0
YResolution: 300.0
ColorSpace: 1
ExifImageWidth: 860
ExifImageHeight: 537
PS C:\Users\rushi\OneDrive\Desktop\python> █
```

Using online application analyze the image

Analysis:

- Digest
- ELA**
- Games
- Hidden Pixels
- ICC+
- JPEG %
- Metadata
- Strings
- Source

U U T A Q




Property	Value
Filename	gory_monalisas_feat_free.jpg
Filetime	2023-10-12 13:27:49 GMT
File Type	image/jpeg
Dimensions	860x537
Color Channels	3
Unique Colors	98243
File Size	412,983 bytes
MD5	d2c906f18e6183e890599d7c63c874a7
SHA1	562233e7c09ae68418a71c5c5844bb63c5852fce
SHA256	850f0c93b40052c857357513f9a37e68dead27d9acd076bac43504cdcd267181
First Analyzed	2023-10-12 13:28:36 GMT

ELA

Analysis:

- Digest
- ELA**
- Games
- Hidden Pixels
- ICC+
- JPEG %
- Metadata
- Strings
- Source

U U T A Q



File

File Type	JPEG
File Type Extension	.jpg
MIME Type	image/jpeg
Exif Byte Order	Big-endian (Motorola, MM)
Current IPTC Digest	a15d6d8a2ad9535ed91050b00305bc46
Image Width	860
Image Height	537
Encoding Process	Baseline DCT, Huffman coding
Bits Per Sample	8
Color Components	3
Y Cb Cr Sub Sampling	YCbCr4:4:4 (1 1)

EXIF

Orientation	Horizontal (normal)
X Resolution	300
Y Resolution	300
Resolution Unit	inches
Software	Adobe Photoshop CS5.1 Windows
Modify Date	2014:04:29 19:14:53
Color Space	sRGB
Exif Image Width	860
Exif Image Height	537
Compression	JPEG (old-style)
Thumbnail Offset	314
Thumbnail Length	5374
Thumbnail Image	(Binary data 5374 bytes)

IPTC

Coded Character Set	UTF8
Application Record Version	108

Photoshop

IPTC Digest	a15d6d8a2ad9535ed91050b00305bc46
Displayed Units X	inches
Displayed Units Y	inches
Print Style	Centered
Print Position	0 0
Print Scale	1
Global Angle	120
Global Altitude	30
URL List	
Slices Group Name	Fig1_Mona Lisa versions and perspective
Num Slices	1
Pixel Aspect Ratio	1
Photoshop Thumbnail	(Binary data 5374 bytes)
Has Real Merged Data	Yes
Writer Name	Adobe Photoshop
Reader Name	Adobe Photoshop CS5.1
Photoshop Quality	12
Photoshop Format	Standard

XMP

XMP Toolkit	Adobe XMP Core 5.0-c061 64.140949, 2010/12/07-10:57:01
Create Date	2013:09:26 18:19:46+01:00
Metadata Date	2014:04:29 19:14:53-04:00
Creator Tool	Adobe InDesign CS5.5 (7.5.3)
Format	image/jpeg
Document ID	xmp.did:C9207A58BCCFE311970EA06040AD7E4D
Instance ID	xmp.iid:12641312F4CFE311970EA06040AD7E4D
Original Document ID	uuid:41545998-cc5d-40b2-a80c-4893030b80c3
Producer	Adobe PDF Library 9.9
Color Mode	RGB

APP14

DCT Encode Version	100
APP14 Flags 0	[14]
APP14 Flags 1	(none)
Color Transform	YCbCr

Composite

Image Size	860x537
Megapixels	0.462

Extract String

Analysis:

Digest

ELA

Games

Hidden Pixels

ICC+

JPEG %

Metadata

Strings

Source

U

V

T

F

A

R

Q



JPEG SOI

JPEG APP1: Exif

0x00000005: 4Exif

0x0000007e: Adobe Photoshop CS5.1 Windows

0x0000009c: 2014:04:29 19:14:53

0x00000140: Adobe_CM

0x0000014e: Adobe

0x00000265: b34r

0x000002ad: 7GWgw

0x000002d4: AQAq"

0x00000305: dEU8te

0x0000032f: 7GWgw

0x00000360: l6A"

0x00000437: 3ljoh

▲

▲

▼

▼

Conclusion

The quest to design and develop a tool for digital forensic analysis of images is driven by the growing significance of visual data in the digital era. As images proliferate across the digital landscape, they have become crucial components of investigations, legal proceedings, and cybersecurity efforts. The traditional methods and tools employed for digital forensics often fall short in efficiently handling the vast and diverse realm of image-based evidence.

The realization of this project addresses this critical gap by aiming to create a robust and specialized tool for image forensics. By doing so, it not only enhances the capabilities of digital forensics experts but also provides valuable support to law enforcement agencies and cybersecurity professionals in their efforts to investigate, analyze, and extract meaningful insights from the complex and rapidly evolving world of digital images.

Reference

1. https://www.researchgate.net/publication/337982446_DESIGN_AND_IMPLEMENTATION_OF_DIGITAL_IMAGE_TOOL_FOR_FORENSIC_ANALYSIS_ANALYSIS
2. <https://www.python.org/doc/>
3. S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," in Advances in Neural Information Processing Systems, 2015.
4. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). "You Only Look Once: Unified, Real-Time Object Detection." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
5. Simonyan, K., & Zisserman, A. (2014). "Very Deep Convolutional Networks for Large-Scale Image Recognition." In International Conference on Learning Representations (ICLR).
6. Kingma, D. P., & Ba, J. (2014). "Adam: A Method for Stochastic Optimization." In International Conference on Learning Representations (ICLR).
7. Abid, A., Awan, A., & Abbas, S. (2018). "Image classification using deep features for content-based image retrieval." Multimedia Tools and Applications, 77(13), 16771-16789.
8. OpenCV. (n.d.). "OpenCV Documentation." <https://docs.opencv.org>. Accessed on [Date Accessed].
9. TensorFlow. (n.d.). "TensorFlow Documentation." <https://www.tensorflow.org>. Accessed on [Date Accessed].