METADATA DETECTION AND REMOVAL OF SECURE IMAGE SHARING USING ML

Under the Supervision of
MRS. NEETU PILLAI

A Project Proposal by
MINHAJ SHAMEER AHAMED
21587119

INTRODUCTION

·Digital image sharing introduces privacy risks due to embedded metadata.

·Metadata (EXIF, IPTC, XMP) can expose sensitive details like geolocation, timestamps, and device information.

·A machine learning–based tool will be developed to detect and remove metadata, ensuring secure image sharing.

LEARNING FROM EXISTING RESEARCHES

ARTICLE TITLE	ISSUES ADDRESSED	METHODOLOGY	LIMITATIONS/ FUTURE RESEARCH
"Privacy-Preserving Metadata Removing using AI" (2019)	Risks from metadata in shared images.	AI-based detection and removal	Needs real-time processing optimization.
"Adaptive Metadata Scrubbing for Social Media" (2020)	Privacy breaches on social media platforms.	Adaptive ML algorithms	Scalability to large datasets.
"Secure Image Sharing Framework with ML" (2020)	Secure sharing of images with metadata risks.	ML integration with encryption techniques	Users' adoption and interface usability.
"Deep Learning for Metadata Analysis" (2019)	Identifying and analysing hidden metadata.	Deep learning models	Limited dataset diversity.

IDENTIFIED RESEARCH GAP

- Inconsistent Metadata Formats and Limited Detection Scope.
- Lack of Standardized Evaluation and Real-Time Processing.
- Privacy, Security, and Social Media Integration Challenges.

AIMS & OBJECTIVES

AIMS

The project aims to develop a machine learning-based tool to detect and remove metadata from digital images, enhancing user privacy during online sharing. It will allow users to sanitize images without compromising quality while promoting awareness of digital privacy.

OBJECTIVES

- Investigate types of image metadata.
- Analyse platform metadata management.
- Develop a metadata removal tool.
- Evaluate tool performance and efficiency.

PROPOSED SOLUTION METHODOLOGY

- Comprehensive Metadata Detection
- Robust Security Measures
- Performance Evaluation and User Feedback Integration
- Compliance with Data Privacy Regulations
- Advanced Machine Learning Algorithms
- User-Centric Design
- Real-Time Processing and Scalability
- Cross-Platform Compatibility

IMPLEMENTATION PROCEDURE

PHASE 1: RESEARCH & PLANNING

- Finalize project topic and requirements.
- Conduct a detailed literature review and identify research gaps.
- Define system architecture and select appropriate tools (Python, TensorFlow, PyTorch).

PHASE 2: DEVELOPMENT & INTEGRATION

- Develop and train the machine learning model for metadata detection and removal.
- Integrate the model into a mobile application with a user-friendly interface.
- Optimize the tool for real-time performance.

PHASE 3: TESTING & DEVELOPMENT

- Conduct usability testing and collect feedback.
- Perform security and performance evaluations.
- Finalize the application and deploy it on mobile platforms (Android/iOS).

ALGORITHM IMPLEMENTATION

1. **Convolutional Neural Network (CNN)**
   - Detects patterns in image data to identify hidden metadata.
   - Provides high detection accuracy for various metadata types.
2. **Random Forest Classifier**
   - Classifies structured metadata (EXIF, IPTC, XMP).
   - Enhances detection accuracy by handling diverse metadata formats.
3. **Autoencoder for Metadata Removal**
   - Reconstructs images without metadata while maintaining image quality.
   - Ensures no loss of visual data during metadata removal.
4. **Natural Language Processing (NLP)**
   - Detects and sanitizes text-based metadata (captions, tags).
   - Provides comprehensive metadata cleansing.

EVALUATION METRICS

ACCURACY

Measurement of how effectively the tool detects and removes metadata.

SECURITY

Assessment of the tool’s ability to prevent metadata leaks and protect user privacy.

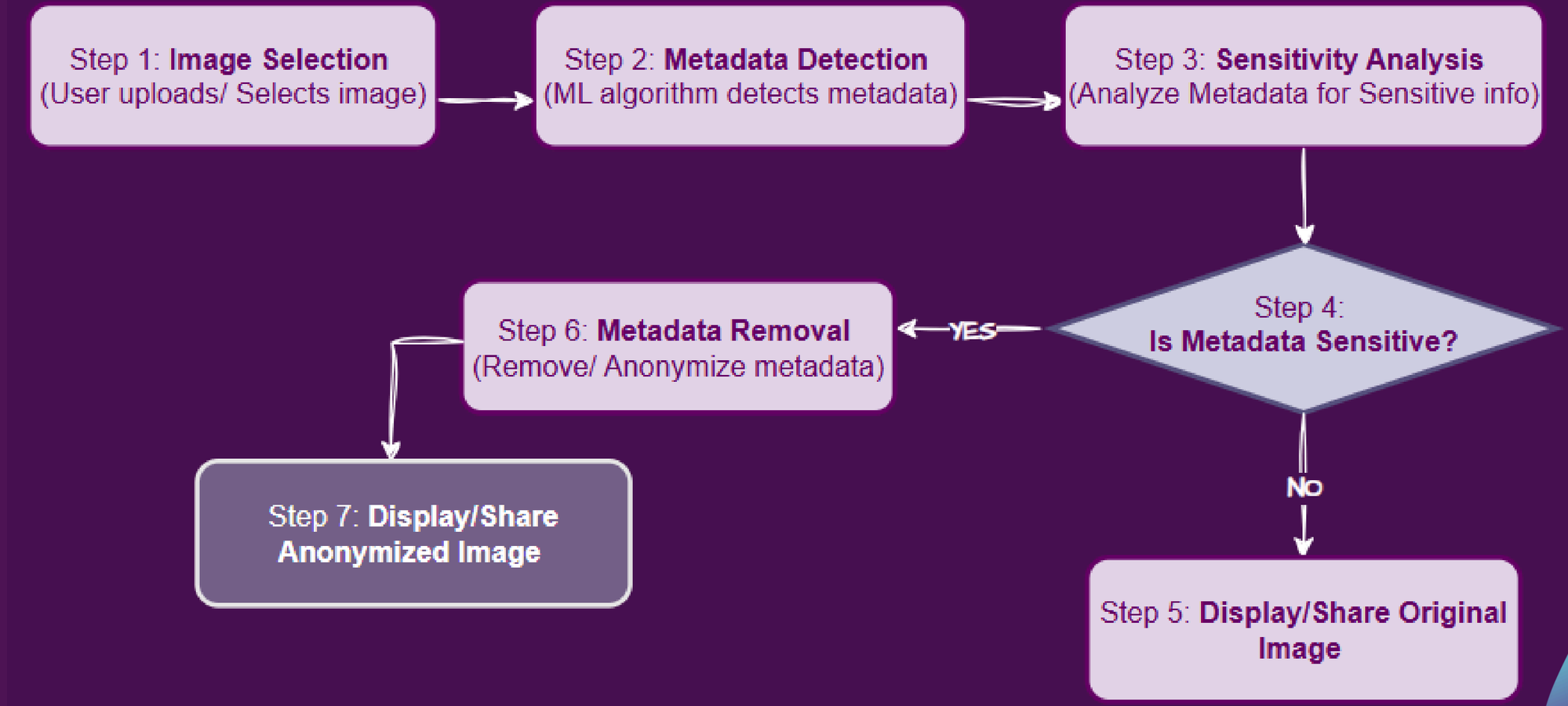
USABILITY

User feedback on ease of use, interface design, and overall user experience.

EFFICIENCY

Processing speed and resource consumption during metadata detection and removal.

PROPOSED MODEL



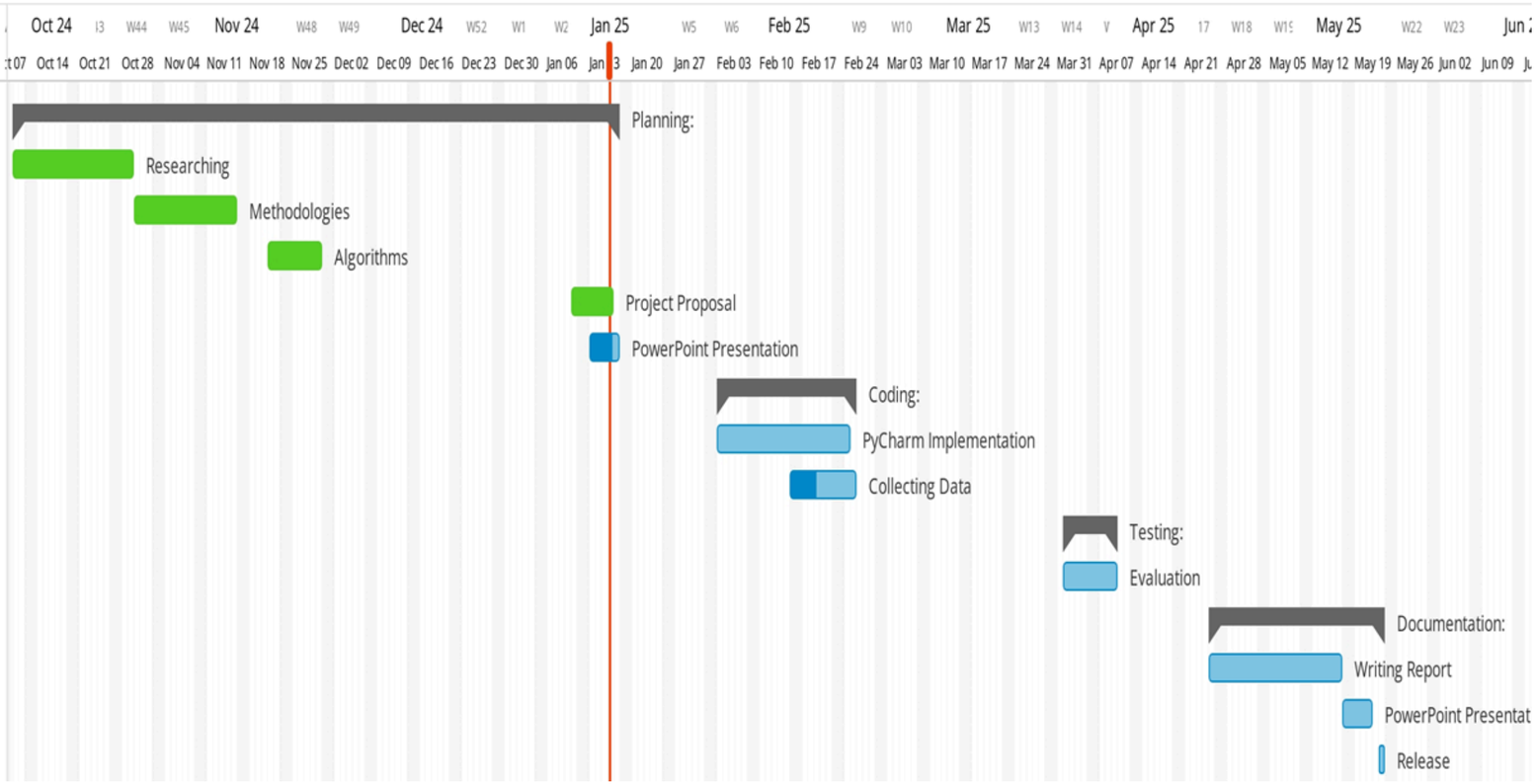
GANTT CHART- PROJECT PLAN & SCHEDULE

metadata

Read-only view, generated on 16 Jan 2025

Instagantt

	ACTIVITIES	START	DUE	STATUS
	Planning:	10/Oct	17/Jan	
1	Researching	10/Oct	29/Oct	Finished
2	Methodologies	30/Oct	15/Nov	Finished
3	Algorithms	21/Nov	29/Nov	Finished
4	Project Proposal	10/Jan	16/Jan	Finished
5	PowerPoint Presentation	13/Jan	17/Jan	Started
	Coding:	03/Feb	25/Feb	
7	PyCharm Implementation	03/Feb	24/Feb	Not started
8	Collecting Data	15/Feb	25/Feb	Started
	Testing:	01/Apr	09/Apr	
10	Evaluation	01/Apr	09/Apr	Not started
	Documentation:	25/Apr	23/May	
12	Writing Report	25/Apr	16/May	Not started
13	PowerPoint Presentation	17/May	21/May	Not started
14	Release	23/May	23/May	Not started



BIBLIOGRAPHY

A. Cohen, N. Nissim, and Y. Elovici, "MalJPEG: Machine Learning Based Solution for the Detection of Malicious JPEG Images," in IEEE Access, vol. 8, pp. 19997-20011, 2020, doi: 10.1109/ACCESS.2020.2969022.

M. Khader and M. Karam, "Assessing the Effectiveness of Masking and Encryption in Safeguarding the Identity of Social Media Publishers from Advanced Metadata Analysis," Data, vol. 8, p. 105, 2023, doi: 10.3390/data8060105.

S. Gayathri and S. Gowri, "Securing Medical Image Privacy in Cloud Using Deep Learning Network," Journal of Cloud Computing, vol. 12, no. 1, p. 40, 2023, doi:10.1186/s13677-023-00422-w.

W. Ma, T. Zhou, J. Qin, X. Xiang, Y. Tan, and Z. Cai, "A Privacy-Preserving Content-Based Image Retrieval Method Based on Deep Learning in Cloud Computing," Expert Systems with Applications, vol. 203, p. 117508, 2022.

N. P. Shetty et al., "Protecting Your Online Persona: A Preferential Selective Encryption Approach for Enhanced Privacy in Tweets, Images, Memes, and Metadata," in IEEE Access, vol. 12, pp. 86403-86424, 2024, doi: 10.1109/ACCESS.2024.3415663.

Gilsang Yoo, Dongeun Sun, Kigon Lyu and Hyeoncheol Kim, "Data hiding technique for digital images tracing system on the Web," 2012 6th International Conference on New Trends in Information Science, Service Science and Data Mining (ISSDM2012), Taipei, Taiwan, 2012, pp. 293-296.

K. P. Arjun et al., "PROvacy: Protecting image privacy in social networking sites using reversible data hiding," 2016 10th International Conference on Intelligent Systems and Control (ISCO), Coimbatore, India, 2016, pp. 1-4, doi: 10.1109/ISCO.2016.7726913.

J. Lepsoy, S. Kim, D. Atnafu and H. J. Kim, "Metadata protection scheme for JPEG privacy & security using hierarchical and group-based models," 2015 5th International Conference on Information & Communication Technology and Accessibility (ICTA), Marrakech, Morocco, 2015, pp. 1-5, doi: 10.1109/ICTA.2015.7426905.

B. Toevs, "Processing of Metadata on Multimedia Using ExifTool: A Programming Approach in Python," 2015 Annual Global Online Conference on Information and Computer Technology (GOCICT), Louisville, KY, USA, 2015, pp. 26-30, doi: 10.1109/GOCICT.2015.14.

Thank you

by MINHAJ SHAMEER
BSc CYBERSECURITY L6
21587119