Image Forensic:Metadata Analysis

Ramya Ajay CB.EN.P2CYS22004
Mohammed Tousif CB.EN.P2CYS22008

TIFAC-CORE in Cyber Security
Amrita Vishwa Vidyapeetham
Coimbatore, India

July 5, 2023

Overview

Introduction

- Metadata for a digital forensic investigator is a unique way to know something or everything that is fused around the actual data. It can be visualized as a cover layer closely surrounding a piece of evidence completely or partially at all times.
- When referring to digital images, metadata can be considered both internal and external to the file containing an image. Internal metadata are usually contained in the EXIF tags defined in for digital image formats and External metadata are represented by the name of the file that OSNs use to store images on their technological infrastructure.
- Here we will analyze the metadata of different image formats.

Existing Methodology

- For image metadata analysis, we use a combination of tools to extract and analyze metadata from images.
- ExifTool, a powerful command-line tool, is commonly used for extracting metadata from various image formats. It provides the ability to extract a wide range of metadata fields, including camera information, timestamps, GPS coordinates, exposure settings, and more.
- But it doesn't give a comprehensive analysis of image metadata rather than listing it.

Proposed Methodology

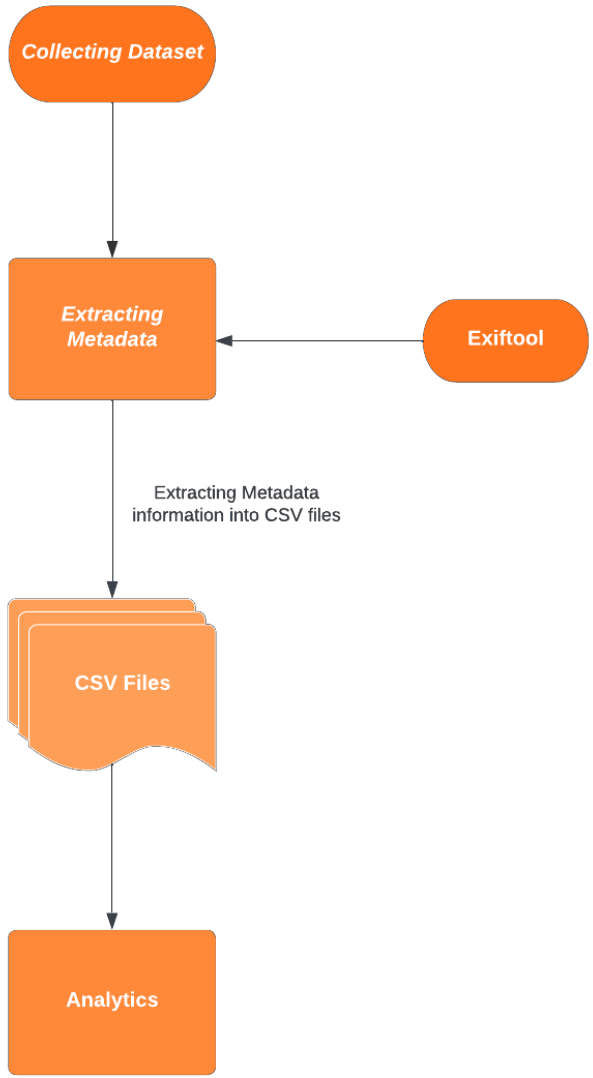
- In our proposed methodology, we primarily focus on raster-based images for metadata analysis. Our dataset consists of a diverse collection of images in various formats to ensure a representative sample for analysis. Below fig,shows the distribution of images in our dataset across different formats.

Image Types	Count
JPG	330
PNG	430
BMP	14
GIF	24
TIFF	196
EPS	24
TGA	26
PCX	57
NEF	11
ORF	15
HEIC	8
DNG	7

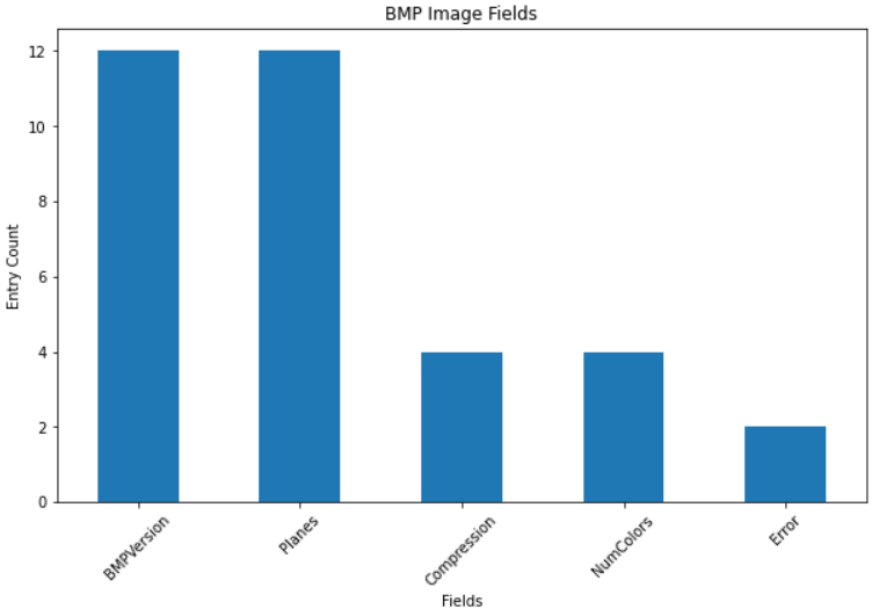
- Figure 2 displays the count of metadata fields extracted from the images in our dataset.

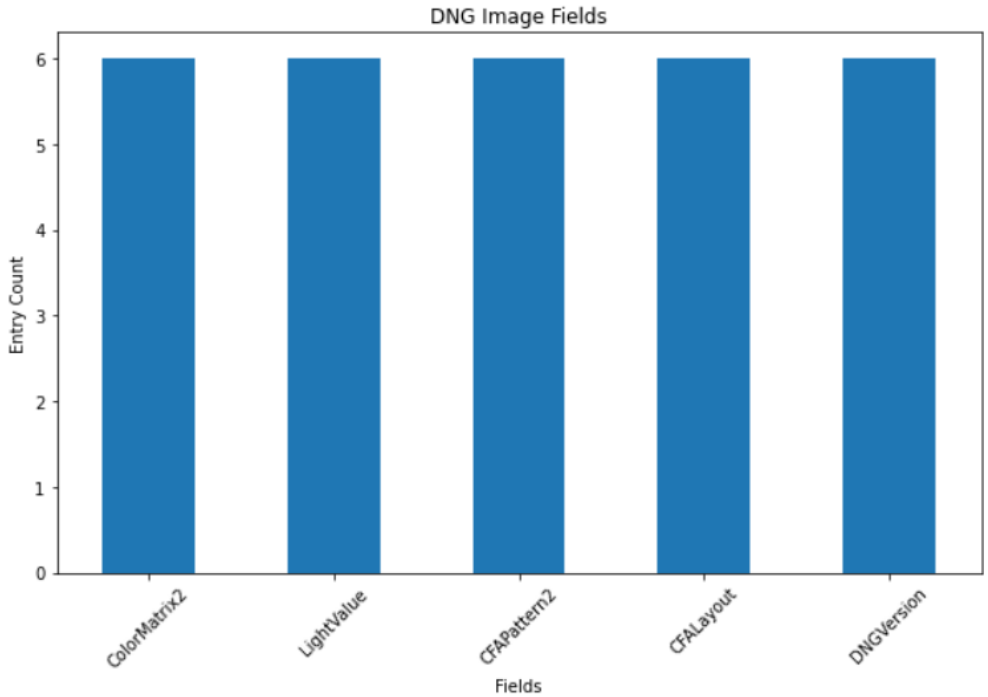
Image Types	Fields Count
JPG	2378
PNG	253
BMP	26
GIF	36
TIFF	722
EPS	324
TGA	14
PCX	29
NEF	494
ORF	292
HEIC	173
DNG	313

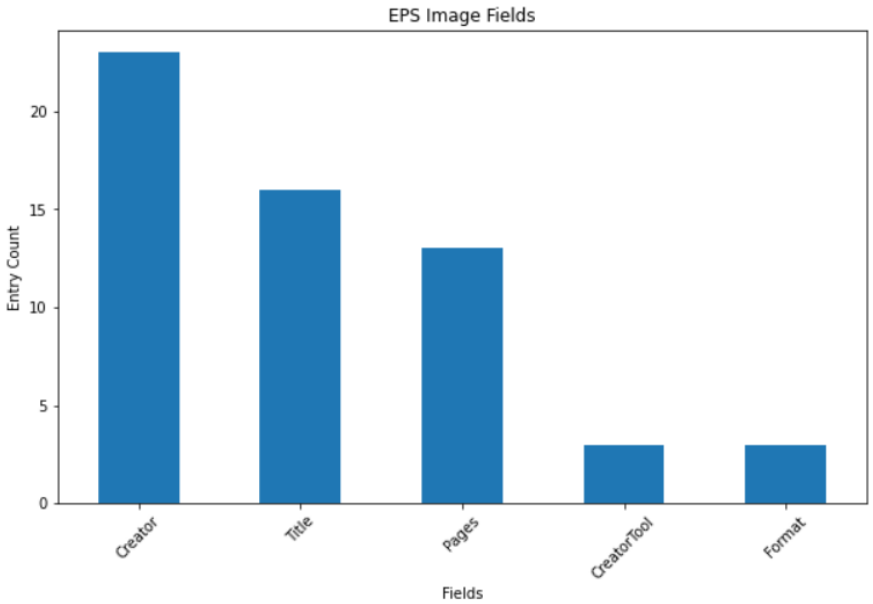
Flowchart

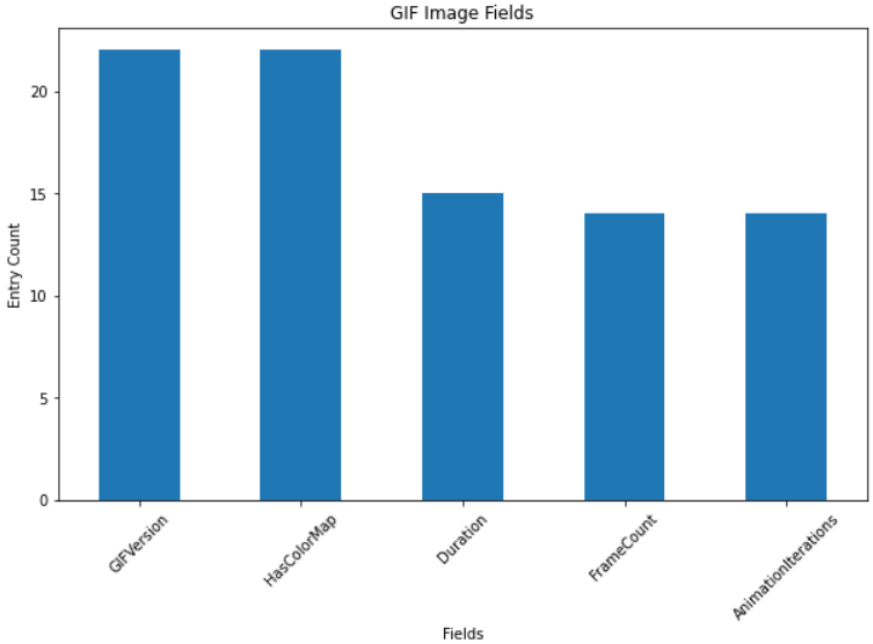

Collecting Dataset
Extracting Metadata
Exiftool
Extracting Metadata information into CSV files
CSV Files
Analytics

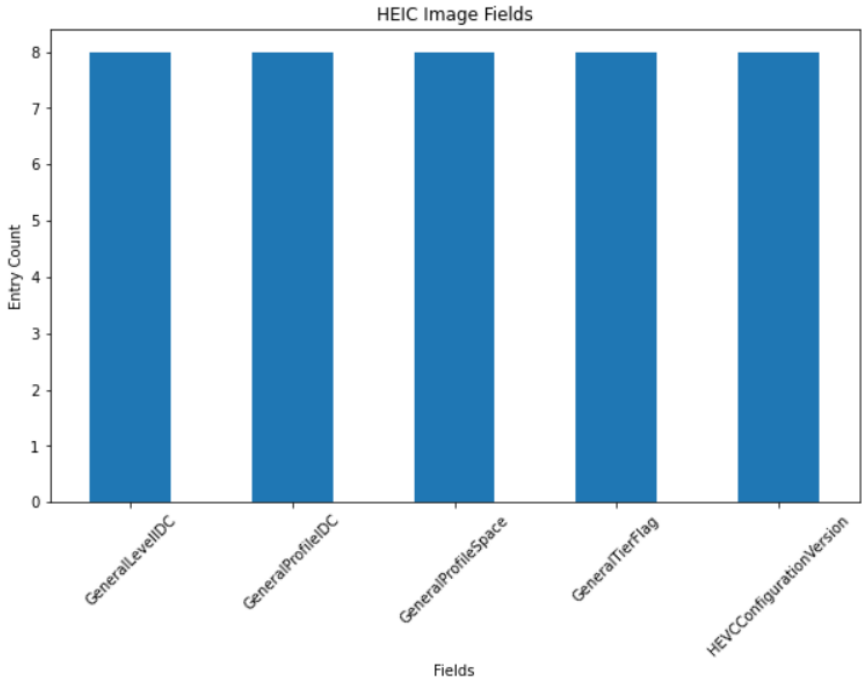
Evaluation

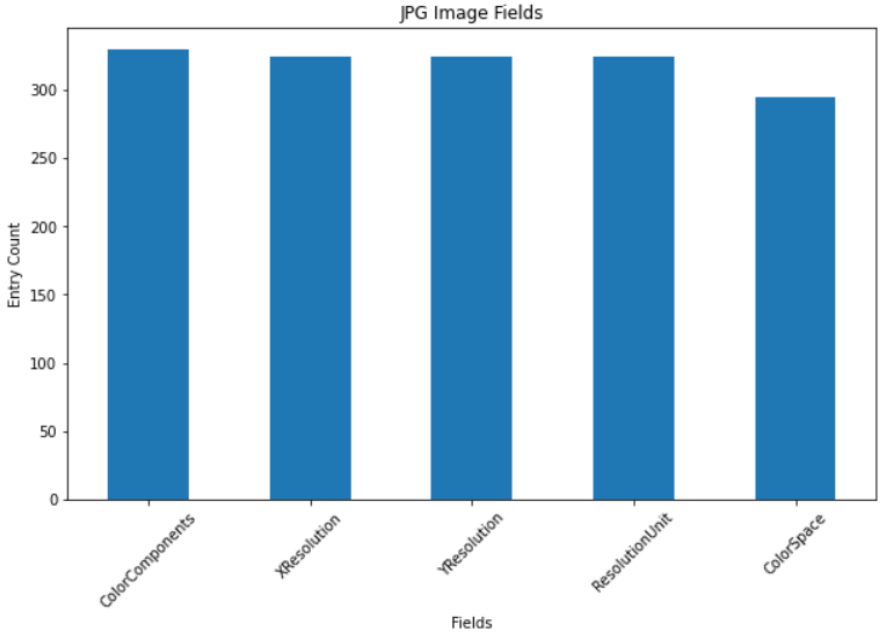

BMP Image Fields
Entry Count
0
2
4
6
8
10
12
BMPVersion
Planes
Compression
NumColors
Error
Fields

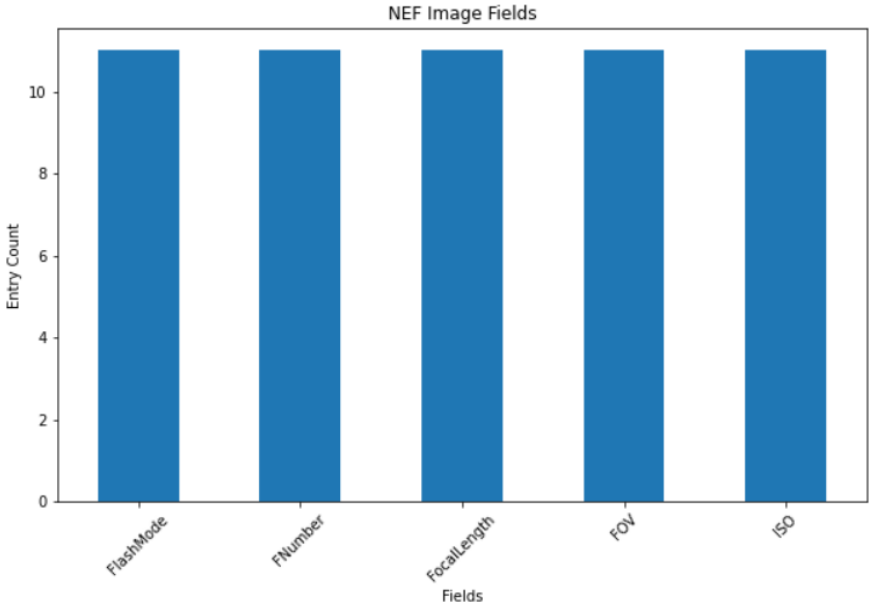

DNG Image Fields
Entry Count
0
1
2
3
4
5
6
ColorMatrix2
LightValue
CFAPattern2
CFALayout
DNGVersion
Fields

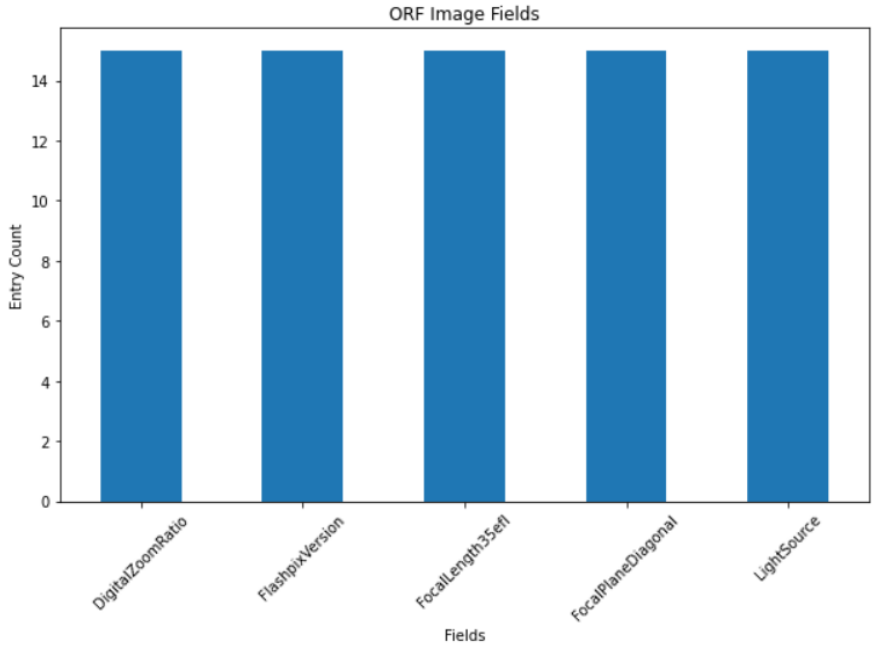

EPS Image Fields
Entry Count
0
5
10
15
20
Creator
Title
Pages
CreatorTool
Format
Fields

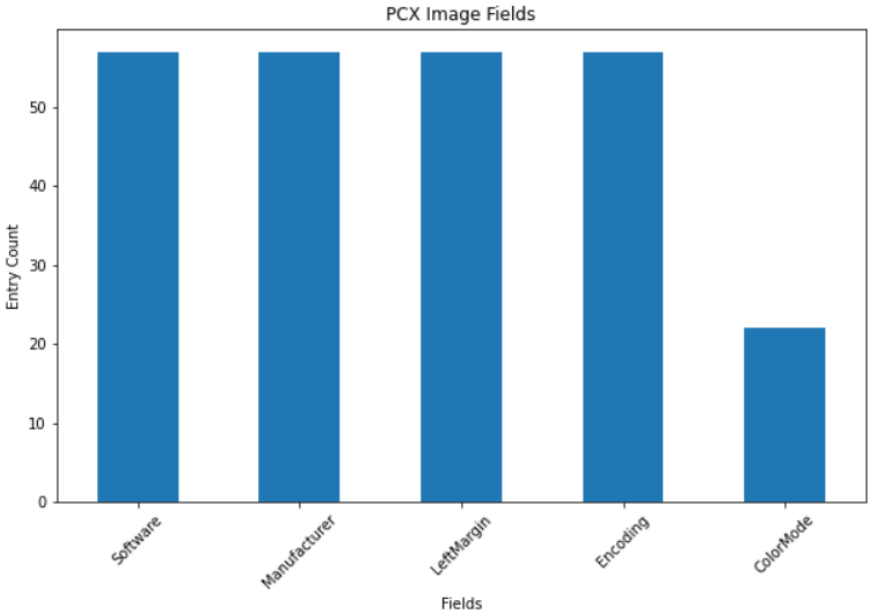

GIF Image Fields
Entry Count
0
5
10
15
20
GIFVersion
HasColorMap
Duration
FrameCount
AnimationIterations
Fields

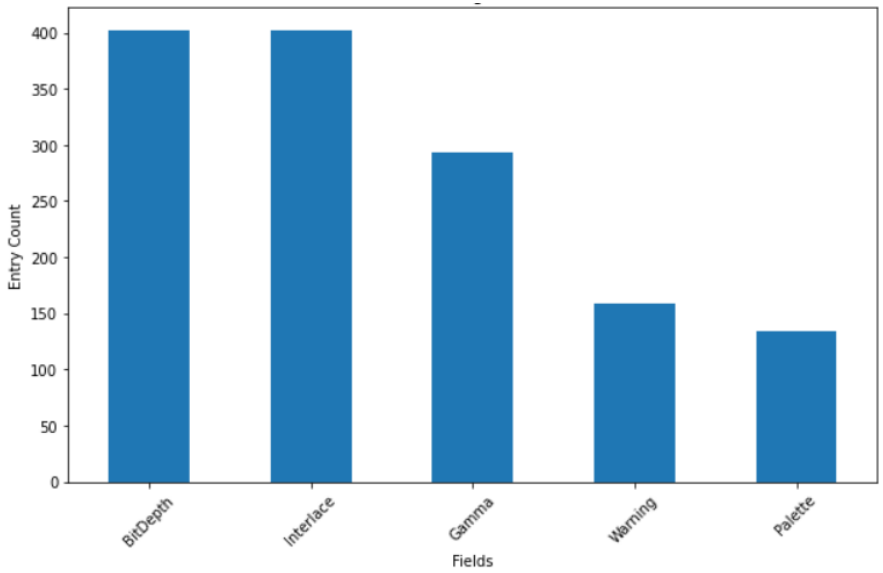

HEIC Image Fields
Entry Count
0
1
2
3
4
5
6
7
8
GeneralLevelIDC
GeneralProfileIDC
GeneralProfileSpace
GeneralTierFlag
HEVCConfigurationVersion
Fields

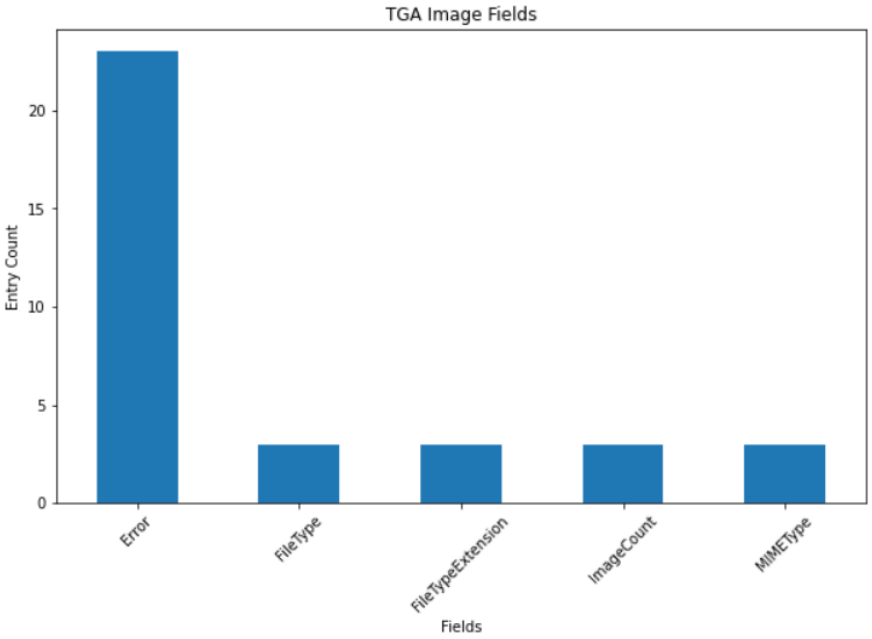

JPG Image Fields
Entry Count
0
50
100
150
200
250
300
ColorComponents
XResolution
YResolution
ResolutionUnit
ColorSpace
Fields

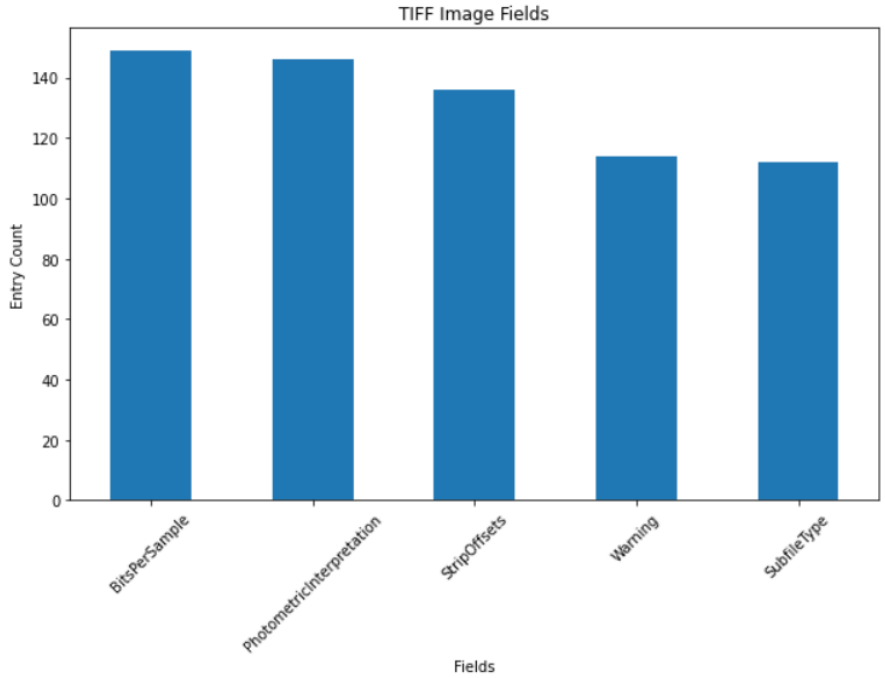

NEF Image Fields
Entry Count
0
2
4
6
8
10
FlashMode
FNumber
FocalLength
FOV
ISO
Fields


ORF Image Fields
Entry Count
0
2
4
6
8
10
12
14
DigitalZoomRatio
FlashpixVersion
FocalLength35efl
FocalPlaneDiagonal
LightSource
Fields


PCX Image Fields
Entry Count
0
10
20
30
40
50
Software
Manufacturer
LeftMargin
Encoding
ColorMode
Fields


Entry Count
0
50
100
150
200
250
300
350
400
BitDepth
Interlace
Gamma
Warning
Palette
Fields


TGA Image Fields
Entry Count
0
5
10
15
20
Error
FileType
FileTypeExtension
ImageCount
MIMEType
Fields


TIFF Image Fields
Entry Count
0
20
40
60
80
100
120
140
BitsPerSample
PhotometricInterpretation
StripOffsets
Warning
SubfileType
Fields

Conclusions and Future work

- In conclusion, ExifTool is a valuable tool that primarily focuses on extracting and presenting the metadata our methodology provides an in-depth analysis and insights of the metadata of different image formats.
- Future work in image metadata analysis can involve expanding the scope of analysis beyond raster images to include vector images. Currently, the dataset used for analysis may be limited to raster image formats such as JPEG, PNG, or TIFF. Including vector image formats such as SVG or AI can provide valuable insights into the metadata characteristics specific to vector graphics. We can also automate the image metadata analysis process by developing advanced algorithms or machine learning models.

References

- Mohan, Ashok Kumar, et al. "Holistic Analytics of Digital Artifacts: Unique Metadata Association Model." IJDCF vol.13, no.5 2021: pp.78-100. http://doi.org/10.4018/IJDCF.20210901.oa5
- exiftool.org "ExifTool by Phil Harvey"
- https://github.com/drewnoakes/metadata-extractor-images "Metadata-Extractor-Images"
- 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.