SIMULATORE METADATI

**Autore**: Vincenzo Faraci

Data: 23/07/2024

Sommario

[1. Introduzione 3](#_Toc172670679)

[1.1. Scopo del progetto 3](#_Toc172670680)

[1.2. Step Descrittivi 3](#_Toc172670681)

[2. Dataset 3](#_Toc172670682)

[2.1. Dati 3](#_Toc172670683)

[2.2. Correlazioni Significative tra Attributi Numerici 4](#_Toc172670684)

[2.3. Strumenti utilizzati 4](#_Toc172670685)

[3. Applicazione Metadati 4](#_Toc172670686)

[3.1. Applicazione compressione 4](#_Toc172670687)

[3.2. Applicazione thumbnail 5](#_Toc172670688)

[3.3. Applicazione metadati 5](#_Toc172670689)

# Introduzione

## Scopo del progetto

Il progetto consiste nello studio dei metadati di dispositivi mobili, in particolare iPhone (modelli studiati: iPhone 14, iPhone 14 Pro, iPhone 14 Pro Max, iPhone 15). L'attività include lo studio dei metadati con conseguente implementazione di algoritmi atti all’applicazione di set di metadati a qualsiasi tipo di immagine.

## Descrizione dati Exif

I dati EXIF (Exchangeable Image File Format) costituiscono un insieme di metadati incorporati automaticamente nelle fotografie digitali o nei file audio durante la loro creazione mediante dispositivi quali fotocamere digitali, smartphone o altri strumenti di registrazione. Questi metadati forniscono una vasta gamma di informazioni tecniche e contestuali riguardanti l'immagine o il file audio.

Contenuto dei dati Exif

1. Informazioni sul Dispositivo di Registrazione:
   * Marca e modello della fotocamera o del dispositivo utilizzato;
   * Numero di serie del dispositivo;
   * Versione del firmware installato;
2. Parametri di scatto:
   * Data e ora in cui la fotografia è stata scattata;
   * Apertura del diaframma (valore f);
   * Velocità dell'otturatore;
   * ISO (sensibilità alla luce);
   * Lunghezza focale dell'obiettivo;
   * Modalità di esposizione;
   * Bilanciamento del bianco;
   * Informazioni sull'uso del flash;
3. Dettagli dell’immagine:
   * Risoluzione dell'immagine.
   * Profondità del colore.
   * Tipo di compressione utilizzata.

## Step Descrittivi

Il lavoro è stato svolto seguendo specifici passi:

**Step 1. Raccolta dei dati.**

I dati sono stati acquisiti da immagini scattate tramite dispositivi iPhone, reperite da varie fonti:

* Camera iPhone 15
* [Flickr](https://www.flickr.com/), in particolare scaricando tramite appositi strumenti (si faccia riferimento alla sezione sugli strumenti utilizzati) le immagini dal gruppo [iPhone 14 Pro su Flickr](https://www.flickr.com/groups/iphone14pro).

Dopo aver scaricato le immagini si è proceduto tramite script Python a estrarne i metadati e costruire il data frame da utilizzare per lo studio statistico di quest’ultimi.

**Step 2. Studio statistico**

Attraverso l’utilizzo della libreria Pandas e della libreria PyExifTool è stato ricavato un dataframe, utilizzato per eseguire studi statistici sui metadati estratti.

Dagli studi statistici si è potuta appurare la natura dei vari dati Exif e ciò a cui fanno riferimento, in particolare quelli specifici delle camere iPhone.

# Dataset

## Dati

Dopo aver scelto i modelli da analizzare, attraverso l’utilizzo di specifici strumenti sono state reperite foto scattate da smartphone iPhone, come descritto in precedenza.

Si è ovviamente prima effettuato uno studio generico sulla natura degli exif.

Gli attributi EXIF coprono diverse categorie:

* Informazioni sul file: nome del file, directory, dimensione, date di modifica/accesso/creazione, permessi, tipo di file, ecc.
* Informazioni sulla fotocamera: modello, produttore, obiettivo, ecc.
* Parametri di scatto: velocità dell'otturatore, apertura, ISO, bilanciamento del bianco, lunghezza focale, ecc.
* Coordinate GPS: latitudine, longitudine, altitudine, ecc.
* Data e ora: data e ora di creazione, modifica, ecc.
* Altro: varie altre informazioni tecniche e composite.

Si è poi passati allo studio del dataset costruito su un campione di 500 immagini.

Dallo studio dei dati estratti si è visto che:

I dati EXIF che si ripetono frequentemente con stesso valore sono:

* **LensSpecification:** indica il modello del sensore del dispositivo
* **SensingMethod:** indica il tipo di sensore del dispositivo (One-chip color area sensor, nel caso di iPhone 14 Pro/Pro Max)
* **Software:** indica la versione del software del dispositivo che ha scattato la foto
* **Model e Make:** indicano modello e marca del dispositivo
* **ISO:** indica il valore ISO della fotocamera
* **Flash:** tramite un valore indica lo stato del flash durante lo scatto della foto
* **ApertureValue:** indica il valore di apertura focale fella fotocamera durante lo scatto

In generale tutti i dati inerenti alla modalità di scatto del sensore sono inevitabilmente simili.

## Correlazioni Significative tra Attributi Numerici

* File Size e Image Width/Height:
  + `File:FileSize` ha una correlazione positiva con `File:ImageWidth` e `File:ImageHeight`, indicando che immagini più grandi in termini di dimensioni del file tendono ad avere risoluzioni maggiori.
* Focal Length e Circle of Confusion:
  + `Composite:FocalLength35efl` ha una correlazione negativa significativa con `Composite:CircleOfConfusion`, suggerendo che una lunghezza focale maggiore tende a ridurre la circolo di confusione.
* Focal Length e Hyperfocal Distance:
  + `Composite:FocalLength35efl` ha una forte correlazione positiva con `Composite:HyperfocalDistance`, indicando che una lunghezza focale maggiore aumenta la distanza iperfocale.
* FOV e Focal Length:
  + `Composite:FOV` ha una correlazione negativa molto forte con `Composite:FocalLength35efl`, il che è previsto poiché un campo visivo più ampio è associato a lunghezze focali più corte.

Questi dati risultano particolarmente utili per la costruzione di un dataset EXIF utile per l'analisi statistica.

## Strumenti utilizzati

Lo studio e la realizzazione dello script sono stati possibili grazie a diversi strumenti e librerie python.

* Python: linguaggio utilizzato per lo script
* Gallery-dl: strumento utilizzato per reperire le immagini dal sito Flickr
* PyExifTool: è una libreria che permette l’utilizzo in python del tool ExifTool per la lettura e la scrittura dei dati exif nelle immagini
* Pandas: libreria python utilizzata per la creazione del dataframe, utilizzato poi per lo studio statistico

# Applicazione Metadati

L'applicazione dei metadati avviene fornendo in input un'immagine (o un gruppo di immagini) alla quale viene applicato lo stesso set di dati EXIF di un'immagine selezionata come template. L'immagine di template è scelta casualmente e può essere facilmente sostituita in base alle esigenze dell'utente.

## Applicazione compressione

Prima di inserire i metadati all’interno dell’immagine è stato effettuato un salvataggio della stessa in modo da poter applicare la stessa compressione (e le tabelle di quantizzazione) dell’immagine originale. Questo ha permesso al file di superare i test effettuati tramite appositi strumenti di controllo come Amped e JPEGSnoop

## Applicazione thumbnail

Aggiungendo i metadati all'immagine, verrebbe inclusa anche la miniatura dell'immagine originale. Per garantire la corretta visualizzazione della miniatura associata all'immagine modificata, seguiamo un processo specifico: inizialmente creiamo la miniatura dell'immagine che desideriamo modificare, successivamente applichiamo i metadati e infine reinseriamo la miniatura originale che avevamo precedentemente salvato. In questo modo, i vari sistemi operativi e strumenti dedicati visualizzeranno la miniatura corretta associata all'immagine.

## Applicazione metadati

Dopo aver estratto e studiato i dati estratti dalle immagini reperite, si passa all’ applicazione di un set exif a delle immagini arbitrarie.

Per applicare questi dati sono stati utilizzati diversi strumenti contemporaneamente, tra cui semplici funzioni python e ExifTool.

Si è visto che tramite ExifTool non tutti i dati sono modificabili direttamente, di seguito una lista dei dati modificabili e non, facendo riferimento ad un set di dati exif standard estratto da una foto casuale:

|  |  |
| --- | --- |
| **Metadati Aggiunti** | **Metadati Non Aggiunti** |
| File:FileModifyDate: | SourceFile: |
| File:FileCreateDate: | ExifTool:ExifToolVersion: |
| File:FilePermissions: | File:FileSize: |
| File:ExifByteOrder: | File:ZoneIdentifier: |
| JFIF:ResolutionUnit: | File:FileAccessDate: |
| JFIF:XResolution: | File:FileType: |
| JFIF:YResolution: | File:FileTypeExtension: |
| EXIF:Make: | File:MIMEType: |
| EXIF:Model: | File:ImageWidth: |
| EXIF:Orientation: | File:ImageHeight: |
| EXIF:XResolution: | File:EncodingProcess: |
| EXIF:YResolution: | File:BitsPerSample: |
| EXIF:ResolutionUnit: | File:ColorComponents: |
| EXIF:Software: | File:YCbCrSubSampling: |
| EXIF:ModifyDate: | EXIF:ThumbnailOffset: |
| EXIF:HostComputer: | EXIF:ThumbnailLength: |
| EXIF:YCbCrPositioning: | EXIF:ThumbnailImage: |
| EXIF:ExposureTime: | MakerNotes:RunTimeFlags: |
| EXIF:FNumber: | MakerNotes:RunTimeValue: |
| EXIF:ExposureProgram: | MakerNotes:RunTimeScale: |
| EXIF:ISO: | MakerNotes:RunTimeEpoch: |
| EXIF:ExifVersion: | MakerNotes:SemanticStyle: |
| EXIF:DateTimeOriginal: | MPF:MPFVersion: |
| EXIF:CreateDate: | MPF:NumberOfImages: |
| EXIF:OffsetTime: | MPF:MPImageFlags: |
| EXIF:OffsetTimeOriginal: | MPF:MPImageFormat: |
| EXIF:OffsetTimeDigitized: | MPF:MPImageType: |
| EXIF:ComponentsConfiguration: | MPF:MPImageLength: |
| EXIF:ShutterSpeedValue: | MPF:MPImageStart: |
| EXIF:ApertureValue: | MPF:DependentImage1EntryNumber: |
| EXIF:BrightnessValue: | MPF:DependentImage2EntryNumber: |
| EXIF:ExposureCompensation: | MPF:MPImage2: |
| EXIF:MeteringMode: | ICC\_Profile:ProfileCMMType: |
| EXIF:Flash: | ICC\_Profile:ProfileVersion: |
| EXIF:FocalLength: | ICC\_Profile:ProfileClass: |
| EXIF:SubjectArea: | ICC\_Profile:ColorSpaceData: |
| EXIF:SubSecTimeOriginal: | ICC\_Profile:ProfileConnectionSpace: |
| EXIF:SubSecTimeDigitized: | ICC\_Profile:ProfileDateTime: |
| EXIF:FlashpixVersion: | ICC\_Profile:ProfileFileSignature: |
| EXIF:ColorSpace: | ICC\_Profile:PrimaryPlatform: |
| EXIF:ExifImageWidth: | ICC\_Profile:CMMFlags: |
| EXIF:ExifImageHeight: | ICC\_Profile:DeviceManufacturer: |
| EXIF:SensingMethod: | ICC\_Profile:DeviceModel: |
| EXIF:SceneType: | ICC\_Profile:DeviceAttributes: |
| EXIF:ExposureMode: | ICC\_Profile:RenderingIntent: |
| EXIF:WhiteBalance: | ICC\_Profile:ConnectionSpaceIlluminant: |
| EXIF:FocalLengthIn35mmFormat: | ICC\_Profile:ProfileCreator: |
| EXIF:SceneCaptureType: | ICC\_Profile:ProfileID: |
| EXIF:LensInfo: | ICC\_Profile:ProfileDescription: |
| EXIF:LensMake: | ICC\_Profile:ProfileCopyright: |
| EXIF:LensModel: | ICC\_Profile:MediaWhitePoint: |
| EXIF:CompositeImage: | ICC\_Profile:RedMatrixColumn: |
| EXIF:Compression: | ICC\_Profile:GreenMatrixColumn: |
| MakerNotes:MakerNoteVersion: | ICC\_Profile:BlueMatrixColumn: |
| MakerNotes:AEStable: | ICC\_Profile:RedTRC: |
| MakerNotes:AETarget: | ICC\_Profile:ChromaticAdaptation: |
| MakerNotes:AEAverage: | ICC\_Profile:BlueTRC: |
| MakerNotes:AFStable: | ICC\_Profile:GreenTRC: |
| MakerNotes:AccelerationVector: | Composite:RunTimeSincePowerUp: |
| MakerNotes:FocusDistanceRange: | Composite:Aperture: |
| MakerNotes:ImageCaptureType: | Composite:ImageSize: |
| MakerNotes:LivePhotoVideoIndex: | Composite:Megapixels: |
| MakerNotes:PhotosAppFeatureFlags: | Composite:ScaleFactor35efl: |
| MakerNotes:HDRHeadroom: | Composite:ShutterSpeed |
| MakerNotes:AFPerformance: |  |
| MakerNotes:SignalToNoiseRatio: |  |
| MakerNotes:PhotoIdentifier: |  |
| MakerNotes:ColorTemperature: |  |
| MakerNotes:CameraType: |  |
| MakerNotes:FocusPosition: |  |
| MakerNotes:HDRGain: |  |
| MakerNotes:RunTimeFlags: |  |
| MakerNotes:RunTimeValue: |  |
| MakerNotes:RunTimeScale: |  |
| MakerNotes:RunTimeEpoch: |  |
| MakerNotes:SemanticStyle: |  |

Bisogna notare che:

* Tra i dati non aggiunti vi sono sia i dati che esistono comunque nel file ma non sono modificabili (es: filesize, sourcefile,ecc..), sia i dati che non sono inseribili in alcun modo e che non esistono nell’immagine (es: gli exif del gruppo MPF)
* tra i metadati non aggiunti ci sono quelli appartenenti al tag Icc\_Profile, questo perché ExifTool non ne permette la modifica diretta, ma essi possono essere cambiati in gruppo.

Nel codice è stato implementato un comando di ExifTool che permette questa operazione. Tutti gli altri tag non sono modificabili tramite ExifTool. Per maggiori informazioni consultare la pagina <https://exiftool.org/TagNames/>