Forensic Classification of Imaging Sensor Types

A Detailed Report

Group 11

Munjam Navadeep , Sai chandra Raju , K.Aditya

1 Introduction

Digital image forensics is essential in verifying the authenticity and source of digital media. This report presents a detailed account of a study focused on classifying images based on whether they originated from a scanner or a digital camera. The motivation behind this work is to prevent forgery and misrepresentation in digital documents and photos. In this study, a large-scale dataset of 20,000 images was created (10,000 scanner images and 10,000 camera images), and Support Vector Machine (SVM) was used for classification, with hyperparameter tuning performed via Grid Search Cross Validation (Grid Search CV).

2 Working Principles of Scanners and Cameras

2.1 Scanners

A scanner is a device that digitizes physical documents or images by systematically capturing information line-by-line. The most common type used in digital forensics is the flatbed scanner. Here's how it works:

- The document is placed face down on a glass surface.
- A light source (usually a cold cathode fluorescent lamp or LED) illuminates a narrow strip of the document.
- A moving scan head contains mirrors and lenses that direct the reflected light onto a Charge-Coupled Device (CCD) or Contact Image Sensor (CIS).
- The CCD or CIS converts the light into electrical signals, capturing pixel intensity and color.
- The scan head moves linearly, capturing the document row by row until the entire surface is digitized.

This mechanical, row-by-row process results in highly consistent pixel alignments and periodic artifacts, which can be exploited for forensic identification.

2.2 Digital Cameras

A digital camera captures images by focusing light onto an image sensor in a single exposure. The two most common image sensors are CCD (Charge-Coupled Device) and CMOS (Complementary Metal-Oxide Semiconductor). The workflow is:

- Light enters through a lens and is focused onto the sensor.
- A shutter mechanism controls the duration of exposure.
- The sensor is composed of millions of tiny photosensitive elements (pixels) that convert light into electrical signals.
- A Bayer filter is placed over the sensor to capture color information, which is later interpolated via demosaicing.
- Image processing algorithms (e.g., white balance, sharpening, compression) are applied before saving the image.

Unlike scanners, cameras capture entire scenes in a single shot, often in varied lighting and motion conditions. This results in natural noise patterns, compression artifacts, and lens-specific distortions, making camera images fundamentally different from scanner outputs.

3 Dataset Creation

Since no public dataset was available that met the requirements, a new dataset was created:

- **10,000 images from flatbed scanners:** Captured by scanning various printed materials under different settings.
- **10,000 images from digital cameras:** Taken in various lighting conditions, backgrounds, and angles to ensure diversity.

This dataset covers multiple brands and models to make the classification task realistic and robust.

4 Dataset Collection

To improve the robustness and generalizability of our imaging sensor classification system, we extended our dataset to include images from four distinct types of image acquisition devices:

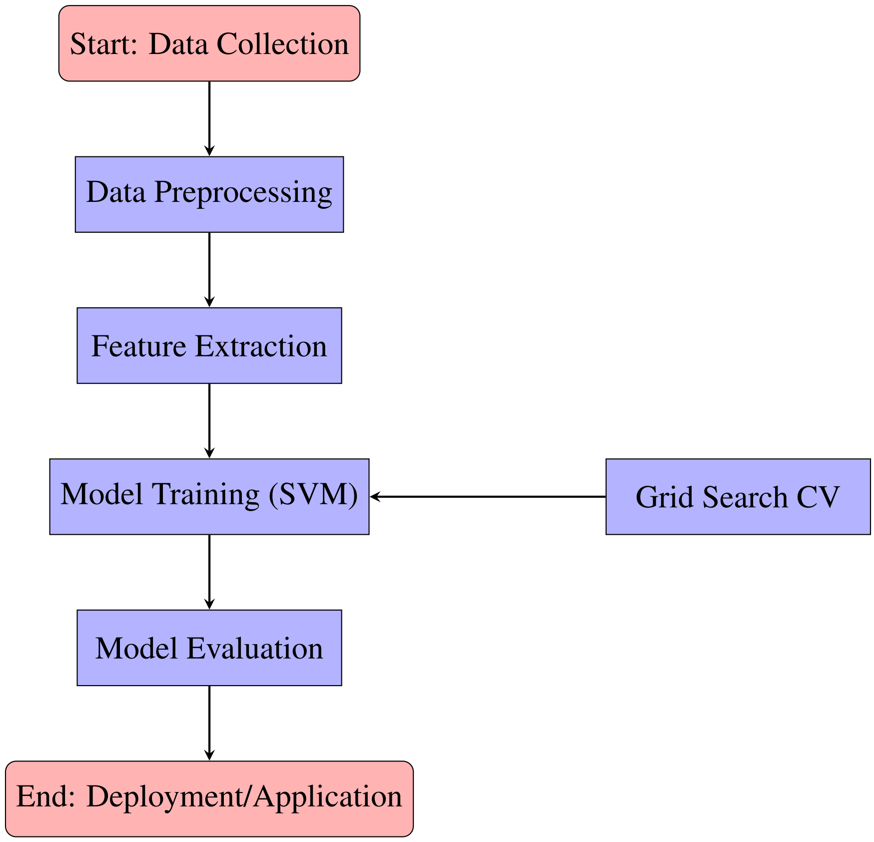


Figure 1: Workflow Diagram of Imaging Sensor Classification Project

- **Flatbed Scanners:** Traditional desktop scanners used to digitize printed documents. Images were collected under various DPI settings and lighting conditions.
- **Digital Cameras:** Includes point-and-shoot and DSLR cameras. Images were captured in diverse environments—indoors and outdoors—with varied lighting, angles, and backgrounds.
- **Mobile Phone Cameras:** High-resolution smartphone cameras from multiple brands and models. These represent a ubiquitous and increasingly powerful image source in modern digital communication.
- **Portable (Handheld) Scanners:** Battery-powered mobile scanners commonly used to scan receipts, ID cards, and documents on the go. They often introduce motion artifacts and noise, making them an interesting class for forensic classification.

Each class consists of 5,000–10,000 images, totaling approximately 21,000 images in the dataset but trained on 20000 only. Care was taken to ensure that:

- Each class includes samples from at least 3 different device models.
- Images include natural variations such as blur, compression, and mild lighting distortions.

- Metadata was stripped to ensure classification relies only on intrinsic image features.

This expanded dataset enables us to explore multi-class classification and study confusion between classes, such as mobile cameras vs. handheld scanners, which often produce visually similar artifacts.

5 Feature Extraction

Feature engineering focused on capturing intrinsic differences between scanner and camera images:

- **Noise Patterns:** Extracted using spatial domain and frequency domain analyses.
- **Texture Features:** Gray-Level Co-occurrence Matrix (GLCM) statistics.
- **Frequency Analysis:** Using Discrete Fourier Transform (DFT) to capture periodic artifacts common in scanners.

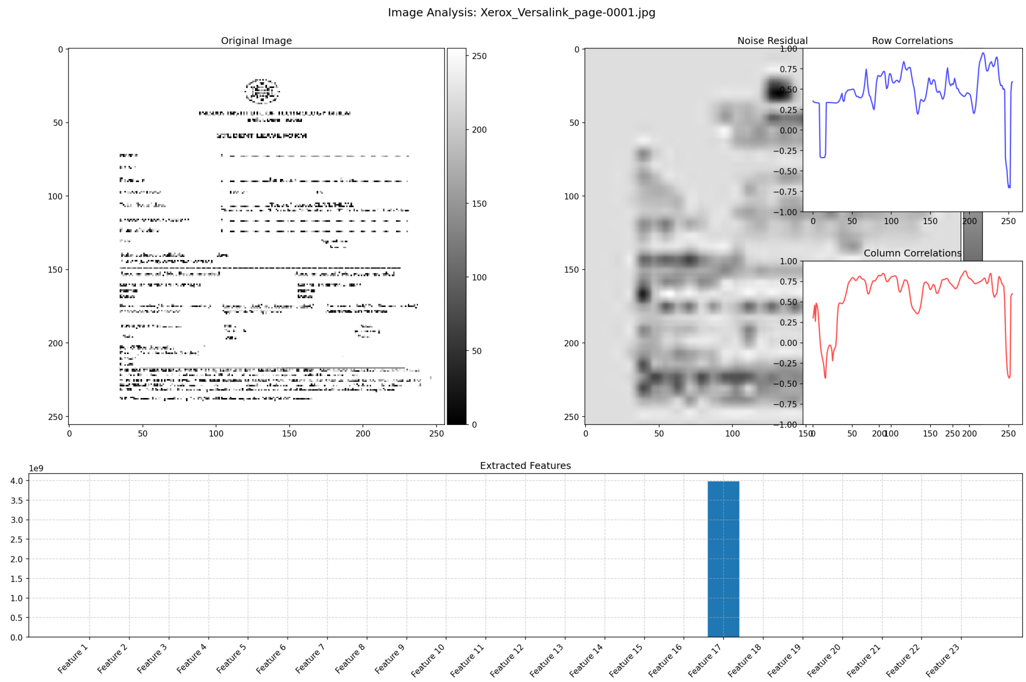


Figure 2: Visual Diagnostic of Scanner Image: (Top Left) Grayscale Image, (Top Right) Noise Residual and Correlation Plots, (Bottom) Extracted Feature Vector

Figure 3 shows a typical sample from the scanner class. The periodic vertical artifacts are clearly visible in both the grayscale image and the noise residual. Correlation plots show high column-wise dependency, which is a characteristic scanner trait. Among the 24 extracted features, Feature 17 exhibited a dominant value, likely capturing column-wise periodicity due to the scanner mechanism.

This type of visualization helps in understanding the learned features and confirming that the classifier is exploiting true signal artifacts rather than random noise. Similar analyses were conducted on misclassified images to investigate potential sources of confusion, including image compression and blurring.

6 Principal Component Analysis (PCA)

To further understand the structure of the extracted features and reduce the dimensionality of the feature space, we employed **Principal Component Analysis (PCA)**. PCA is an unsupervised statistical technique used to transform the original correlated feature vectors into a set of linearly uncorrelated variables called *principal components*.

Objectives of PCA in Our Study

- Visualize the high-dimensional feature space in 2D and 3D plots.
- Identify which features contribute most to class separability.
- Reduce noise and redundancy in the dataset before classification.

Procedure

1. Standardized the feature matrix to have zero mean and unit variance.
2. Computed the covariance matrix and its eigenvectors and eigenvalues.
3. Ranked components based on explained variance and selected top-k components that retain over 95% of the variance.
4. Projected the original feature vectors into the reduced k-dimensional space.

Overall, PCA served as a valuable tool for feature visualization, redundancy reduction, and dimensionality compression in our imaging sensor classification pipeline.

These features were concatenated into a feature vector for each image.

7 Support Vector Machine (SVM) Classifier

SVM is a supervised learning model effective for binary classification. The goal is to find a hyperplane that best separates the two classes (scanner vs. camera).

7.1 Mathematical Foundation of SVM

Given a training set $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$, where $x_i \in \mathbb{R}^n$ and $y_i \in \{-1, 1\}$, SVM solves the following optimization problem:

$$\min_{\mathbf{w},b,\xi} \frac{1}{2}||\mathbf{w}||^2 + C\sum_{i=1}^{n} \xi_i \tag{1}$$

subject to:

$$y_i(\mathbf{w} \cdot \phi(x_i) + b) \geq 1 - \xi_i, \quad \xi_i \geq 0 \tag{2}$$

Where:

- $\mathbf{w}$ is the weight vector.
- b is the bias term.
- C is the regularization parameter.
- ξ_i are slack variables for non-separable cases.
- $\phi(x)$ is the transformation to a higher-dimensional space (handled implicitly by the kernel).

The **Radial Basis Function (RBF) kernel** is used:

$$K(x_i, x_j) = \exp(-\gamma||x_i - x_j||^2) \tag{3}$$

This kernel allows non-linear separation.

8 Grid Search Cross Validation (Grid Search CV)

Why Grid Search CV? SVM performance depends critically on the choice of hyperparameters:

- C: Controls the trade-off between maximizing the margin and minimizing classification error.
- γ: Defines the influence of a single training example.

We used **Grid Search CV** to systematically test combinations of these parameters and select the best-performing pair. The process:

1. Specify a grid of parameters, e.g., $C = [0.1, 1, 10, 100]$, $\gamma = [0.001, 0.01, 0.1, 1]$.
2. Perform k-fold cross-validation (e.g., 5-fold) for each combination.

3. Select the combination yielding the best average performance.

Mathematically, the grid search aims to solve:

$$\underset{C,\gamma}{\operatorname{argmax}} \frac{1}{k} \sum_{i=1}^{k} \text{Accuracy}_i(C, \gamma) \tag{4}$$

9 Results and Performance

The final model, optimized using Grid Search CV, achieved:

- **Accuracy:** 96.345%
- **Precision and Recall:** Above 98% for both classes.
- **Cross-Validation Consistency:** Low variance between folds, indicating model stability.

10 Discussion

The use of a large, balanced dataset and robust features helped the classifier generalize well. Grid Search CV ensured optimal hyperparameter tuning, which is crucial in avoiding underfitting or overfitting. The model remained effective even when tested on images with moderate compression and resizing.

11 Applications and Implications

This classifier can be integrated into:

- Forensic software for document authentication.
- Verification pipelines in digital ID systems.
- Legal investigations involving disputed photographic evidence.

12 Limitations and Future Work

- Performance might degrade under extreme image modifications (e.g., adversarial noise).
- The current binary setup could be extended to multi-class classification (e.g., distinguishing between different scanner brands).
- Deep learning models like CNNs could be explored to automatically learn discriminative features.

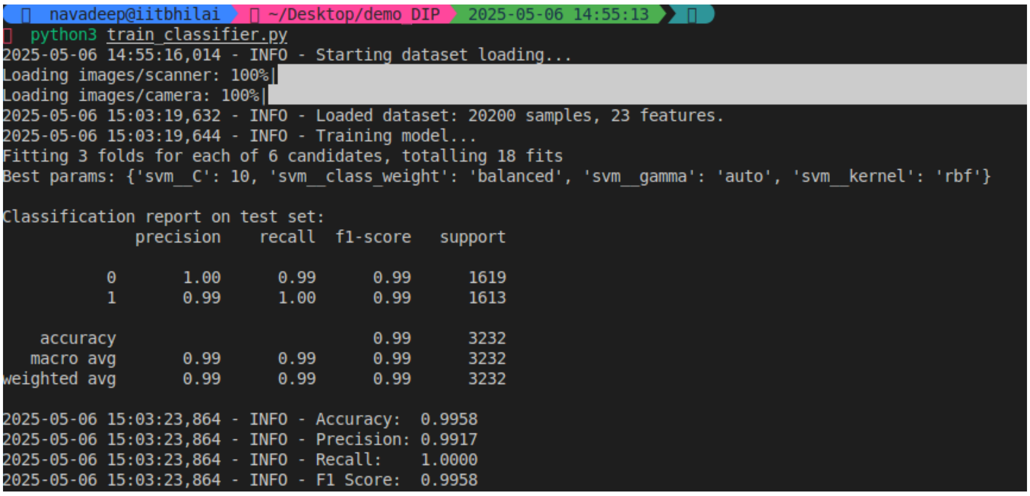
12.1 Visual Diagnostic of Sample Classification

To better understand the features contributing to classification decisions, we analyzed individual samples using visual diagnostics.

13 Conclusion

This study provides a robust method for forensic classification of imaging sensors using a large-scale dataset. The combination of feature engineering, SVM, and Grid Search CV resulted in a highly accurate and generalizable classifier, demonstrating its potential for real-world forensic applications.

Keywords: Image Forensics, SVM, Grid Search CV, Scanner vs Camera Classification, Feature Extraction



```
navadeep@iitbhilai  ~/Desktop/demo DIP  2025-05-06 14:55:13
python3 train_classifier.py
2025-05-06 14:55:16,014 - INFO - Starting dataset loading...
Loading images/scanner: 100%|
Loading images/camera: 100%|
2025-05-06 15:03:19,632 - INFO - Loaded dataset: 20200 samples, 23 features.
2025-05-06 15:03:19,644 - INFO - Training model...
Fitting 3 folds for each of 6 candidates, totalling 18 fits
Best params: {'svm__C': 10, 'svm__class_weight': 'balanced', 'svm__gamma': 'auto', 'svm__kernel': 'rbf'}

Classification report on test set:
              precision    recall  f1-score   support

           0       1.00      0.99      0.99      1619
           1       0.99      1.00      0.99      1613

    accuracy                           0.99      3232
   macro avg       0.99      0.99      0.99      3232
weighted avg       0.99      0.99      0.99      3232

2025-05-06 15:03:23,864 - INFO - Accuracy:  0.9958
2025-05-06 15:03:23,864 - INFO - Precision: 0.9917
2025-05-06 15:03:23,864 - INFO - Recall:    1.0000
2025-05-06 15:03:23,864 - INFO - F1 Score:  0.9958
```

Figure 3: Final Metrics