

CSCE 5222 Feature Engineering – Project Report

Gabor Filter Bank Features for Real vs Altered Fingerprint Classification (SOCOFing)

Group 7: Amir Naderian, Alireza Mohammadshafie

1) Problem Statement

We evaluate whether a compact Gabor filter-bank representation can reliably distinguish authentic (“Real”) from tampered (“Altered”) fingerprints in the SOCOFing dataset. Because fingerprints are composed of oriented ridge-valley textures, orientation- and frequency-selective Gabor filters are a natural choice for feature extraction. A robust Real-vs-Altered detector is important for security (spoof and identity-fraud prevention), operational quality control during enrollment/verification, and forensic triage. Our goal is an accurate, lightweight solution—faster and easier to deploy than heavy deep models—by summarizing Gabor responses into a small feature vector and classifying with standard learners.

2) Data

- Dataset: SOCOFing (fingerprint images)
- Classes: Real (authentic prints) and Altered (includes Easy/Medium/Hard subfolders).
- Counts detected:
 - Real: 6000 images
 - Altered: 49270 images
 - Total: 55270 images
- Preprocessing: convert to grayscale (handling RGB/RGBA/single-channel), resize to 128×128 for consistency.



3) Method

3.1 Preprocessing

- Read image
- Convert to grayscale
- Resize to 128×128 (anti-aliased, reflect mode)
- Normalize to $[0, 1]$

3.2 Gabor Filter Bank and Features

- Orientations (degrees): 0, 45, 90, 135, 180
- Frequencies: 0.1, 0.2, 0.3
- For each (orientation, frequency) pair: compute complex Gabor response; use magnitude; summarize with mean and standard deviation.
- Final feature vector per image: $2 \times (5 \text{ orientations}) \times (3 \text{ frequencies}) = 30 \text{ features.}$

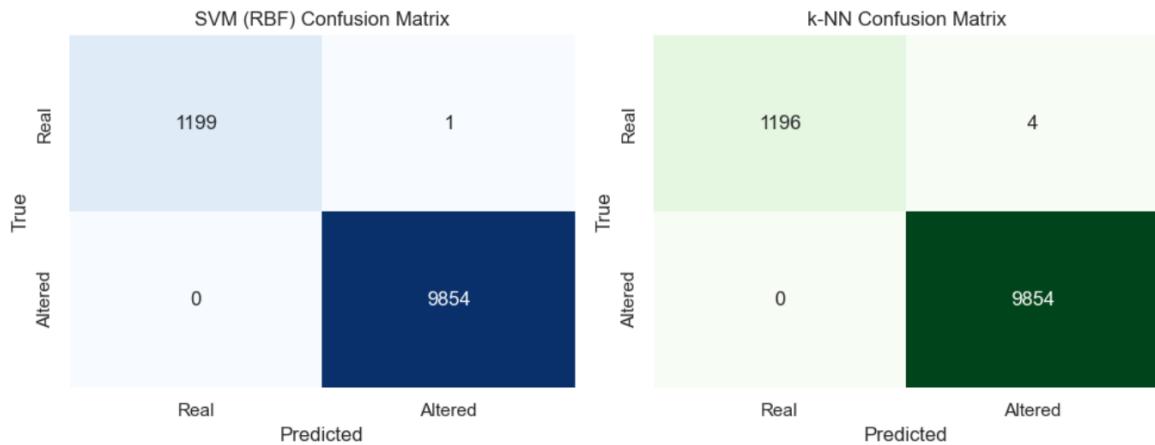
3.3 Splitting, Scaling, and Models

- Split: 20% test; 20% validation from the remaining pool (stratified).
 - Train: 33162 samples; Val: 11054; Test: 11054
- Scaling: StandardScaler fit on train features; applied to val/test.
- Classifiers:
 - SVM (RBF), class_weight=balanced
 - Grid: $C \in \{0.1, 1, 10, 100\}$; gamma $\in \{\text{scale}, 0.01, 0.001\}$
 - CV: Stratified 5-fold
 - k-NN
 - Grid: n_neighbors $\in \{3, 5, 7, 9\}$; metric $\in \{\text{euclidean}, \text{manhattan}\}$
 - CV: Stratified 5-fold

4) Evaluation

- Metrics: Accuracy and Precision (Altered treated as positive class).
- Confusion matrices and classification reports generated on the test set.

5) Results



SVM (RBF) Classification Report:

	precision	recall	f1-score	support
Real	1.00	1.00	1.00	1200
Altered	1.00	1.00	1.00	9854
accuracy			1.00	11054
macro avg	1.00	1.00	1.00	11054
weighted avg	1.00	1.00	1.00	11054

k-NN Classification Report:

	precision	recall	f1-score	support
Real	1.00	1.00	1.00	1200
Altered	1.00	1.00	1.00	9854
accuracy			1.00	11054
macro avg	1.00	1.00	1.00	11054
weighted avg	1.00	1.00	1.00	11054