Identification of Fake Currency Notes

1. Definition

1.1. Project Overview

Counterfeit currency poses a significant challenge for major economies worldwide, as it undermines the value of legitimate money and contributes to inflation. Instances of economic warfare through counterfeit operations have been observed historically, such as the notable counterfeit scandal in Hungary in 1926. The prevalence of counterfeit bills in circulation is alarming, with the US Department of Treasury estimating approximately $70 million of counterfeit currency in the United States alone.

To combat this issue, innovative techniques and countermeasures are necessary to accurately identify fake currency notes. Leveraging the power of computers and technology, machine learning can play a crucial role in developing tools for this purpose. By training machine learning models on patterns and features extracted from genuine and counterfeit currency images, we can teach the models to recognize distinguishing characteristics.

The goal of this project is to create a robust classification model capable of accurately identifying counterfeit banknotes with minimal classification errors. This supervised learning model will be trained using a labeled dataset sourced from the UCI Machine Learning Repository. The dataset consists of images captured from both genuine and forged banknote-like specimens, with features extracted using the Wavelet Transform technique.

By leveraging image processing and pattern recognition techniques, the trained model will learn to differentiate between genuine and counterfeit banknotes based on the identified decision boundaries. This model can provide a convenient and reliable solution to help individuals and institutions detect counterfeit currency and mitigate its adverse effects on the economy.

Overall, the project aims to utilize machine learning techniques to develop an effective tool for the identification of counterfeit banknotes, contributing to the ongoing efforts in combating this pervasive issue.

1.Variance of wavelet transformed image (continuous)

2. Skewness of wavelet transformed image (continuous)

3. Kurtosis of wavelet transformed image (continuous)

4. Entropy of image (continuous)

5. Class of the currency (integer)

1.2. Problem Statement

The problem statement is formally defined as: “Given a dataset containing features extracted after wavelet transformation of images of currency notes, use the features available in the dataset to design a supervised learning model which can identify whether a currency note is real or fake.”

2. Analysis

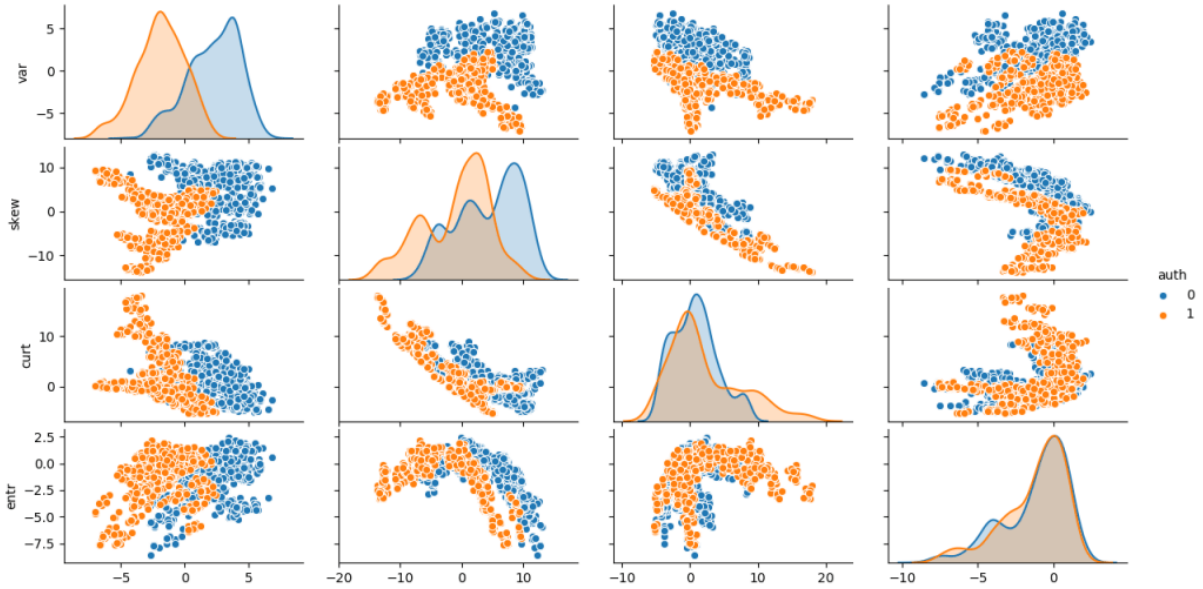
2.1. Data Exploration

The banknote authentication dataset consists of five attributes, viz., variance, skewness, kurtosis, entropy and the class. This dataset consists of 1372 samples, out of which 762 correspond to fake notes, while 610 correspond to real notes
A brief description of the dataset, including parameters like mean, min, max for each column is given in table.

	variance	skewness	kurtosis	entropy	class
count	1372.000000	1372.000000	1372.000000	1372.000000	1372.000000
mean	0.433735	1.922353	1.397627	-1.191657	0.444606
std	2.842763	5.869047	4.310030	2.101013	0.497103
min	-7.042100	-13.773100	-5.286100	-8.548200	0.000000
25%	-1.773000	-1.708200	-1.574975	-2.413450	0.000000
50%	0.496180	2.319650	0.616630	-0.586650	0.000000
75%	2.821475	6.814625	3.179250	0.394810	1.000000
max	6.824800	12.951600	17.927400	2.449500	1.000000

Exploratory Visualization

The distribution of both variance and skewness appears to be quite different for the two target characteristics, while kurtosis and entropy appear to be more similar.
There are clear linear and nonlinear trends in the input features.
Some characteristics seem to be correlated.
Some features seem to separate genuine and fake banknotes quite well.



The logistic regression model achieved an accuracy of 98.36%. And not only that, when our fake currency detection model predicted that a banknote was real, it was correct 100% of the time.