International Journal of Social Science and Economic Research

ISSN: 2455-8834

Volume:10, Issue:10 "October 2025"

Analyzing Banknote Authenticity Through Neural Classification Models And Imaging Statistics

Akaash Sachdeva[1] and Guillermo Goldsztein[2]

[1]Langley High School McLean, VA

[2]Georgia Institute of Technology, Atlanta, Georgia

DOI: 10.46609/IJSSER.2025.v10i10.047 URL: https://doi.org/10.46609/IJSSER.2025.v10i10.047

Received: 23 September 2025 / Accepted: 21 October 2025 / Published: 30 October 2025

ABSTRACT

An increasing challenge to international financial companies, the rise of counterfeit money erodes institutional trust and threatens economic stability. Modern, data-driven solutions are required since traditional fraud detection techniques have not been able to keep up with the $70 million worth of counterfeit bills that are still in circulation in the United States alone. This study investigates the statistical analysis of scanned photos using supervised machine learning to identify counterfeit banknotes. Wavelet Transform was used to process each banknote image using a publicly available dataset in order to extract four important features: entropy, kurtosis, variance, and skewness. These characteristics offer a measurable foundation for model training by representing underlying distribution and texture patterns. A 3-layer neural network, a 4-layer neural network, and logistic regression were the three models that were assessed. Every model was trained to identify whether the notes were authentic or counterfeit, and it consistently performed well. Near-perfect precision and recall were achieved by the 4-layer neural network, which yielded the highest overall accuracy (98.2%). However, the logistic regression model, despite its simplicity, yielded an impressive 94.7% accuracy, making it an efficient and interpretable alternative for real-world deployment. With ramifications for integration into ATMs, retail systems, and mobile verification applications, these findings show the effectiveness of very simple machine learning models in fraud detection activities. In an increasingly digitized economy, this study provides a scalable and accurate framework for preventing financial fraud by fusing binary classification algorithms with statistical imaging approaches.

I. Introduction

Banknote fraud is the creation and use of counterfeit currency, where fraudulent money is passed off as legitimate. This illegal process is detrimental to a country's financial system, undermining

ISSN: 2455-8834

Volume:10, Issue:10 "October 2025"

the state of currency by causing unnecessary inflation and depriving legitimate businesses of revenue and taxes, affecting overall economic growth. Moreover, it has recently, and surprisingly, shown increasing relevance, where the world has recently witnessed $22 million in banknotes seized by the United States Secret Service in 2023, and a 19% increase in the use of forged notes and coins reported by the EU [1]. The U.S. still has up to $70 million in counterfeit cash in circulation, and it has caused an increased priority to combat counterfeit crimes, with ambitions for businesses and the federal government to identify fake bills, surrender them, and alert authorities. The introduction of machine learning to combat this problem implements significant benefits to the process, allowing for a more accurate, quick, and reliable method of banknote fraud detection. Similar actions have innovated other areas of federal security, such as in surveillance detection routes [2].

II. Supervised Learning And The Dataset

A. Details of The Dataset

The dataset contained data extracted from images taken from both genuine and forged banknotes. An industrial camera was used, and it collected 400x400 pixel images with 660 dpi (dots per inch) resolution [3]. The dataset utilizes Wavelet Transform, a mathematical tool that decomposes an image into different frequency components, where features vary over different scales. By utilizing this tool on the images, each image can be broken down into a set of data points, and apply machine learning to these measurements to determine a difference in values between real and fake notes. There are a total of 1,372 data points for the set, with four features for training. The ratio between determined features and points of data is fair, as there are many more points for training than aspects that affect the result.

B. Features

The dataset's features , which are the values that describe each data point for the model to use in training, are variance, skewness, kurtosis, and entropy.

Variance measures the dispersion of pixel values in an image, reflecting the level of contrast or texture complexity. In machine learning, a higher value of variance can likely indicate intricate patterns on the associated image, which led me to hypothesize that a larger variance would be typical in the genuine banknotes, due to intricate details in real bills that the government implements to differentiate them from fake ones [4].

Skewness quantifies the asymmetry of pixel value distribution, which can help determine whether the data has a bias towards higher or lower intensities. Skewness can be used to distinguish differences in skewness between the counterfeit money and the genuine ones.

ISSN: 2455-8834

Volume:10, Issue:10 "October 2025"

Kurtosis measures the “tailedness” of the pixel value distribution. This indicates the presence of outliers or extreme values in the image.

Entropy describes the randomness or complexity of the image’s pixel values, with higher entropy indicating more complexity.

C. Data

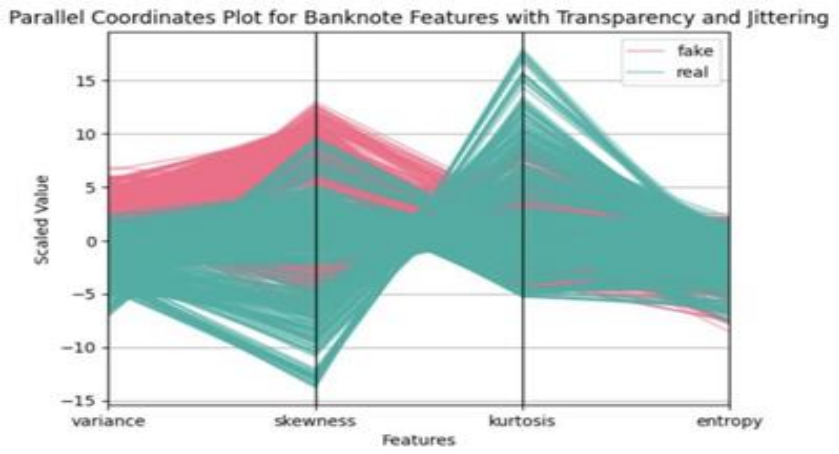
The collection of data is illustrated below in Table 1, where two examples of data from the table are shown.

TABLE I. Data of two randomly selected examples from the dataset

	variance	skewness	kurtosis	entropy	class
Ex. 1	1.0637	3.6957	-4.1594	-1.9379	1
Ex. 2	1.5940	4.7055	1.3758	0.08188	0

This set of data is partly representative of the dataset’s overall quantitative landscape; as seen in the comparison between the first row and the second, the row with the class feature of ‘1’ has a smaller variance, smaller skewness, smaller kurtosis, and a smaller entropy. Although not the case for all the data points, or even the overwhelming majority, the real bills, with a ‘1’ representative of ‘true’, had a significantly smaller value for variance, a slightly smaller value for skewness and entropy, and a slightly larger value for kurtosis overall. Such trends will be demonstrated in the upcoming table (Figure 1), and the only misleading piece of information from this small comparison of a large dataset of 1,372 values was the notably smaller value in the kurtosis in the real bill.

Fig. 1. Plotting of features for all bills in the dataset



ISSN: 2455-8834

Volume:10, Issue:10 "October 2025"

Shown above in Table 2 is the distribution of points from real and counterfeit currency based on the four features. The pink lines represent the fake banknotes, whereas the teal lines represent the real banknotes. As visible in the chart, and much like the previous example, in general the fake banknotes had a much larger variance and skewness than the real ones, disproving my hypothesis. This may be because the material used on actual banknotes may cause less phenomena in the imaging, however there is no concrete answer coming out of this experiment.

On the other hand, the kurtosis of the real banknotes was significantly higher than the fake ones. This shows that using these three will be especially helpful in comparison. Regarding the last feature, the entropy of both seems relatively similar. However, what this image does not show is the larger amount of real bills with higher entropy than fake bills.

Although methods were taken to minimize this inconvenience in viewing the last feature, such as adding transparency to the graph and 'jittering' the values, meaning that they are moved slightly to prevent overlap, it is still difficult to analyze the last one purely based on graphs. Nevertheless, it is a helpful column that will assist in further clarifying the model's output.

III. Experiment

A. Classification

In the model, the type of problem faced is binary classification. Because the result is either a "yes" or "no", this situation falls into the field of binary classification since the project outputs a binary value.

B. Logistic Regression

Logistic regression is a statistical model that estimates the probability of an event occurring based on a set of independent variables [5]. To understand the logistic regression used in the experiment, a basic explanation of the sigmoid function is needed. Though the equation is unimportant for the sake of the process, the graph is displayed below.

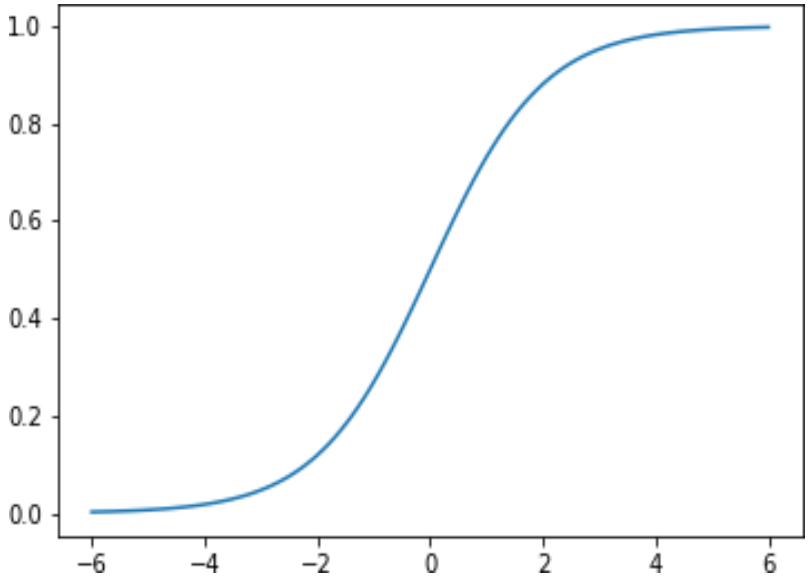
Important attributes are that the function ranges from 0 to 1, exclusive (*(0,1)*), and as x approaches the left of the graph, the graph approaches 0. Conversely, as x approaches the right of the graph, the function approaches 1. Lastly, the center of the graph, at an x the function outputs 0.5

Therefore, the banknote model will take in the 4 features, weigh them in a way to reduce error in the training set, and then output a value from the function. This was the exact functioning of the first model.

ISSN: 2455-8834

Volume:10, Issue:10 "October 2025"

Fig. 2. The Sigmoid function, the first model's method of estimation, graphed, [-6,6]



A. Neural Networks

The next two models utilized neural networks. A neural network is another type of model that aims to mimic the way the human brain makes decisions [6]. In neural networks, there is an input layer, where the model takes in all the features, there are hidden layers in the middle, where the network will further weigh the significance of each feature, and then there is the output layer, where the network will give an answer. Additionally, in each of these layers are nodes, which describe how many sets of weights the network will transfer through each layer. In this situation, since the project is using binary classification, the output layer will have two nodes, representing two categories, and the input layer will have four nodes, representing four features.

IV. Results

As mentioned earlier, three separate models were tested with the data, in hopes of differentiating accuracy of the outputs. However, an important thing to note is that, while three models were tested, the only reasons three were created was because it was the framework of the project. As the results displayed, the first model had an acceptable accuracy, and, being logistic regression, was both precise and concise. The other, more detailed models generated better results; however, the output highlighted from the experiment is the first model, utilizing logistic regression.

A. Logistic regression model

- *5.3% Validation error*

ISSN: 2455-8834

Volume:10, Issue:10 "October 2025"

- *5.4% Training error*
- *98% precision*
- *98% recall*
- *94.7% accuracy*

B. 3-layer neural network

- *1.9% Validation error*
- *52.4% Training error*
- *98% precision*
- *99% recall*
- *98.1% accuracy*

C. 4-layer neural network

- *1.8% Validation error*
- *1.8% Training error*
- *99% precision*
- *99% recall*
- *98.2% accuracy*

The results indicate significant accuracy, through precision, recall, and an overall minimized error of the data points. More, the jump from the 3-node neural network and the 4-node neural network was noticeably minimal, likely hinting towards an approach state of accuracy for a long-running model. The results, though all satisfactory, point towards a promising model specifically of logistic regression in real- world implementation.

VI. Conclusion

The project illustrates machine learning's use in identifying fake currency, utilizing statistical information obtained from image analysis. Several models were trained and their performance in

ISSN: 2455-8834

Volume:10, Issue:10 "October 2025"

binary classification assessed using a dataset that included values for variance, skewness, kurtosis, and entropy. The logistic regression model offered a straightforward but incredibly accurate answer, with 94.7% accuracy with little validation and training error, whereas more intricate models like neural networks achieved somewhat higher accuracy. These results demonstrate how even simple machine learning models, which provide speed, accuracy, and scalability, may greatly help with fraud detection jobs. Machine learning offers a viable path to improving financial security through automation and data-driven decision-making, since counterfeit detection continues to be a significant problem, particularly in light of recent surges in circulation. Altogether, the invention poses as a free, consistent, implementable solution that incorporates innovative technologies that are assisting similar fields greatly.

V. Applications of Further Use

More real-world applications are made possible by machine learning models' ability to identify counterfeit currency. To quickly confirm the legitimacy of currency, financial institutions can incorporate comparable models into cash counters, automated teller machines (ATMs), and mobile banking applications. Furthermore, these techniques can be included into portable verification tools that law enforcement or shops employ, offering a convenient and effective detection option.

In addition to banknotes, forged documents, false identification, and even counterfeit objects with distinctive visual patterns can be identified using the texture-based image feature analysis method. By directly learning complicated spatial characteristics from photos, future advancements such as combining deep learning with convolutional neural networks (CNNs) could significantly improve accuracy.

A centralized fraud-monitoring network may also be formed by integrating this classification system with real-time data reporting, which would speed up the identification of counterfeit distribution patterns across geographical boundaries. Building more flexible and robust detection systems will require ongoing machine learning solution integration and improvement as financial fraud changes.

ACKNOWLEDGEMENT

This research was conducted as part of the AI: Machine Learning and Python course, under Professor Guillermo Goldsztein, in the School of Mathematics at Georgia Tech.

REFERENCES

[1] Old School Fraud: Counterfeit Cash Makes Global Comeback. (2024, March 15).

ISSN: 2455-8834

Volume:10, Issue:10 "October 2025"

PYMNTS. Retrieved October 4, 2025, from https://www.pymnts.com/news/security-and-risk/2024/old-school-fraud-counterfeit-cash-makes- global-comeback

[2] P., T., K., S., S., P. T., & C., S. M. (2024). Detecting Counterfeit Currency with Image Processing. Journal of Cognitive Human-Computer Interaction. Retrieved October 4, 2025, from https://americaspg.com/article/pdf/2663

[3] DPI resolution: Your complete guide. (n.d.). Adobe. Retrieved October 4, 2025, from https://www.adobe.com/uk/creativecloud/photograph y/discover/ dots-per-inch-dpi-resolution.html?msockid=2d625d63ec7c6b220c1048 a8ed6e6ae9

[4] Bank-Note-Authentication. (2020). Kaggle. Retrieved October 4, 2025, from https://www.kaggle.com/datasets/shanks0465/bankno teauthentication

[5] Lee, F. (n.d.). What is logistic regression? IBM. Retrieved October 4, 2025, from https://www.ibm.com/think/topics/logistic-regression

[6] Lee, F. (n.d.). What is a neural network? IBM. Retrieved October 4, 2025, from https://www.ibm.com/think/topics/neural-networks

