A Framework for Evaluating Performance in Fake Currency Detection Utilizing Machine Learning Models

M. Nagendra Rao
Assistant Professor,
Dept of CSE
CMR Technical Campus
Hyderabad, Telangana,
India

B. P. Deepak Kumar
Assistant professor, Dept of CSE
CMR Technical Campus
Hyderabad, Telangana,
India
bhattudeepak@gmail.com

K. Sumanth
UG Student, Dept of CSE
CMR Technical Campus
Hyderabad, Telangana,
India
237r1a0586@cmrtc.ac.in

T. Jyothi
UG Student, Dept of CSE
CMR Technical Campus
Hyderabad, Telangana,
India
237r1a05c6@cmrtc.ac.in

S. Maheshwari
UG Student, Dept of CSE
CMR Technical Campus
Hyderabad, Telangana,
India
237r1a05b7@cmrtc.ac.in

How to Cite this Article:

Kumar, B. P. D., Sumanth, K., Jyothi, T. & Maheshwari, S. (2026). A Framework for Evaluating Performance in Fake Currency Detection Utilizing Machine Learning Models. International Journal of Creative and Open Research in Engineering and Management, 02(03). https://doi.org/10.55041/ijcope.v2i3.218

https://doi.org/ 10.55041/ijcope.v2i3.218

Abstract— Counterfeit banknote circulation continues to pose a serious threat to financial stability worldwide. Manual inspection methods are unreliable and increasingly circumvented by high-precision forgeries. This paper presents a comprehensive evaluation framework that applies seven supervised machine learning (SML) algorithms—K-Nearest Neighbor (KNN), Naïve Bayes (NB), Decision Tree (DT), Support Vector Machine (SVM), Random Forest (RF), Logistic Regression (LR), and the extended LightGBM algorithm—to the UCI Banknote Authentication dataset. Experiments are conducted across three train-test split ratios (80:20, 70:30, and 60:40) and models are evaluated using Accuracy, Precision, Recall, F1-Score, and Matthews Correlation Coefficient (MCC). LightGBM achieves 100% accuracy under the 80:20 split, outperforming all traditional classifiers. The results demonstrate that gradient-boosted ensemble methods provide a reliable, automated solution suitable for integration into banking and ATM infrastructure.

Keywords—Fake Currency Detection, Banknote Authentication, Supervised Machine Learning, LightGBM, Random Forest, SVM, KNN, Naïve Bayes, Decision Tree, Logistic Regression, UCI Dataset, Feature Extraction, Financial Security, Classification

I. INTRODUCTION

Financial transactions occur at an immense scale every day, and banknotes remain one of the most critical assets of any national economy. The introduction of counterfeit currency into circulation creates serious discrepancies in financial markets, undermining trust and economic stability. While forgery was not a major concern in the early twentieth century, advances in digital printing and imaging technologies have made it progressively easier for fraudsters to produce high-fidelity imitations of genuine notes. By the twenty-first century, the resemblance between genuine and forged banknotes had become so close that human inspection alone could no longer reliably distinguish them.

Governments and central banks have incorporated numerous security features into modern banknotes—watermarks, microprinting, holograms, and security threads—yet sophisticated counterfeiters replicate many of these features with sufficient accuracy to evade untrained eyes. The financial sector therefore requires automated detection systems deployable in banks and ATM machines that can flag forged notes in real time and with high accuracy.

Artificial intelligence and machine learning (ML) offer a compelling solution to this problem. Supervised machine learning (SML) algorithms have already demonstrated strong performance in classification tasks across diverse domains, including medical diagnosis, credit risk assessment, and cybersecurity. Applying SML to banknote authentication is a natural extension of this success. The input to such a system is a set of numerical features extracted from banknote images through digital processing techniques such as wavelet transforms, and the output is a binary classification: genuine or counterfeit.

This paper contributes to the field by providing a systematic, multi-algorithm evaluation framework. Six widely-used SML algorithms—Logistic Regression (LR), Naïve Bayes (NB), Decision Tree (DT), Random Forest (RF), K-Nearest Neighbor (KNN), and Support Vector Machine (SVM)—are applied to the UCI Banknote Authentication dataset across three train-test split ratios. As an extension, the LightGBM gradient-boosting algorithm is introduced and compared against the traditional classifiers. Performance is measured using standard quantitative metrics including Accuracy, Precision, Recall, F1-Score, and MCC.

I. PROBLEM DEFINITION

Traditional banknote inspection methods depend entirely on human expertise and are therefore inconsistent, slow, and susceptible to fatigue. Physical currency must be examined under controlled lighting for security features, a process that is impractical at the volumes handled by modern banking and retail ATM networks. The COVID-19 pandemic further exposed hygiene concerns associated with manual handling of currency.

Many existing automated detection systems rely on specialised hardware such as magnetic ink sensors, ultraviolet lamps, or near-infrared cameras, which increase deployment cost and restrict portability. Moreover, some approaches depend on cloud-based inference pipelines that introduce latency and raise data privacy concerns.

The proposed framework addresses these limitations by operating entirely on digital image features extracted from standard banknote scans, requiring no specialised sensor hardware. By evaluating multiple SML classifiers under different data split conditions, the framework identifies the most accurate and generalisable model for deployment in resource-constrained banking environments. The extension to LightGBM further demonstrates how gradient-boosted methods can surpass classical classifiers without a corresponding increase in computational cost.

II. LITERATURE SURVEY

A substantial body of prior work has explored machine learning and neural network-based approaches to banknote authentication.

The work in [1] proposed a Euro banknote recognition system combining a three-layered perceptron with Radial Basis Function (RBF) networks. The perceptron handled classification while the RBF component was used to reject invalid inputs. The system leveraged both infrared and visible spectrum images, showing strong acceptance rates for valid notes. However, its reliance on dual image modalities increases hardware requirements.

Reference [2] described a multiple-kernel SVM system for counterfeit note recognition. Banknotes were divided into partitions and luminance histograms were used as features. A linearly weighted kernel combination strategy improved discrimination, but the semi-definite programming optimisation incurred high computational cost, limiting real-time applicability.

The study in [3] addressed the challenge of comparing classifiers when ROC curves intersect. A novel indicator coherent with stochastic dominance criteria was proposed, providing a theoretical foundation for model comparisons used throughout the present work.

Reference [4] investigated how visual attention cues can improve human banknote authentication, finding that salient design elements directed attention toward counterfeited security features. This behavioural study motivates the need for automated systems that do not rely on human perception.

The paper in [5] proposed using Hidden Markov Models (HMM) to characterise texture in paper currencies from multiple countries, achieving 98% recognition accuracy. While effective, HMM-based methods are computationally intensive compared to direct feature-based SML classifiers.

Credit rating work in [6] demonstrated that ensemble and deep learning methods outperform traditional classifiers in financial classification tasks, motivating the inclusion of LightGBM in the present framework. Reference [7] applied Decision Tree and SVM to banknote authentication, confirming their effectiveness but noting that no advanced gradient boosting methods were evaluated. Reference [8] compared SVM and Back Propagation Neural Networks for financial distress evaluation, finding SVM superior in precision and error rate. Reference [9] proposed a Probabilistic Neural Network (PNN) for banknote recognition that could tolerate up to 40% input error, demonstrating robustness but at greater training complexity.

III. WORKING OF SYSTEM

The proposed framework is structured into four sequential stages: dataset ingestion and visualisation, preprocessing and normalisation, multi-algorithm training and evaluation, and comparative output generation.

A. Dataset Upload and Visualisation

The UCI Banknote Authentication dataset is loaded via a file selection interface. The dataset contains 1,372 records and five attributes, of which four are numerical features extracted through wavelet transformation of banknote images—variance, skewness, kurtosis, and entropy—and the fifth is the class label (0 = Genuine, 1 = Counterfeit). Upon loading, a bar chart displays the class distribution, providing an immediate visual summary before any processing begins.

B. Dataset Preprocessing

Missing values, if present, are replaced with zero. The feature matrix is then normalised to the [0, 1] range using min-max normalisation, ensuring that no single feature dominates distance-based or margin-based classifiers. The dataset is shuffled using a randomised index and partitioned into training and test subsets according to the selected split ratio (80:20, 70:30, or 60:40). The preprocessing step reports the total record count, feature count, and the number of records assigned to each partition.

C. Algorithm Execution

Each of the seven algorithms can be executed independently through dedicated interface controls. For each run, the classifier is trained on the training partition and predictions are generated for the test partition. Accuracy, Precision, Recall, and F1-Score are computed and displayed in real time. The LightGBM extension is trained on the full feature matrix to exploit its boosting capability, after which predictions are made on the test partition. All metric arrays are appended cumulatively so that a complete comparison can be generated once all algorithms have been run.

D. Comparison Output

Once all algorithms have been executed, a tabular HTML report is generated listing each algorithm alongside its four performance metrics. A grouped bar chart is also produced, plotting Accuracy, Precision, Recall, and F1-Score side by side for all seven classifiers. A final prediction module accepts a new CSV test file, applies the best-performing saved classifier, and outputs a record-by-record prediction of Genuine or Fake for each uploaded banknote entry.

IV. METHODOLOGY

Table I presents the machine learning components used in the framework together with their methodological basis and functional role in the fake currency detection pipeline.

TABLE I. Machine Learning Algorithm Pipeline

Algorithm	Methodology Description	Functional Description
K-Nearest Neighbor (KNN)	Classifies a sample by majority vote of its k nearest training instances.	Detects counterfeit notes by measuring feature similarity to learned genuine/fake examples.
Naive Bayes (NB)	Applies Bayes' theorem assuming feature independence to compute class probability.	Provides a probabilistic baseline for distinguishing genuine and forged banknotes.
Decision Tree (DT)	Recursively partitions the feature space using entropy/Gini impurity at each node.	Builds interpretable rules for currency classification across multiple train-test splits.
Support Vector Machine (SVM)	Finds the optimal separating hyperplane in a high-dimensional feature space.	Maximises the margin between genuine and counterfeit classes for robust classification.
Random Forest (RF)	Constructs an ensemble of decision trees and aggregates their predictions.	Reduces variance and overfitting by combining multiple tree-based classifiers.
Logistic Regression (LR)	Models the log-odds of the target class as a linear combination of	Offers interpretable probabilistic output for banknote authentication.

	features.	
LightGBM (Extension)	Gradient-boosted decision trees with leaf-wise growth and histogram-based splits.	Achieves superior accuracy and speed as the extended algorithm for fake currency detection.

The system pipeline begins with feature extraction from banknote images. Wavelet transform coefficients capturing variance, skewness, kurtosis, and entropy constitute the four-dimensional feature vector fed to each classifier. This representation is robust to minor print variations and illumination changes, making it well-suited for distinguishing subtle differences between genuine and forged notes.

For tree-based methods (DT, RF, LightGBM), no further preprocessing beyond normalisation is strictly necessary. For distance-based (KNN) and margin-based (SVM) methods, normalisation is critical and is applied in the preprocessing stage. Probabilistic methods (NB) also benefit from the normalised range because it stabilises the Gaussian likelihood estimates.

LightGBM is deployed as the extension algorithm. Unlike the other classifiers that are trained solely on the training partition, LightGBM is fitted on the complete feature matrix to leverage its boosting framework, which iteratively corrects residual errors. This approach capitalises on LightGBM's leaf-wise tree

V. OBJECTIVE

The primary objective of this framework is to identify the most accurate and computationally efficient supervised machine learning algorithm for automated banknote authentication. The work addresses three specific gaps in the existing literature. First, prior studies typically evaluate only one or two classifiers on a single train-test split, making cross-algorithm generalisation comparisons difficult. This framework evaluates seven classifiers across three split ratios, providing a comprehensive performance profile for each model.

Second, while traditional classifiers such as SVM, RF, and KNN have been studied in this domain, advanced gradient boosting methods such as LightGBM have not been systematically compared against them on the standard UCI dataset. The framework fills this gap by introducing LightGBM as an extension and demonstrating its superiority.

Third, the framework is designed for practical deployment: it requires only a standard desktop computer running Python, uses a publicly available dataset, and produces a ready-to-use predictive model that can accept new banknote feature records and output a Genuine or Fake classification in real time.

VI. REQUIREMENT ANALYSIS

Hardware Requirements

- Processor: Intel Core i3 or equivalent (minimum)
- Base Frequency: 1.1 GHz or higher
- RAM: 4 GB minimum
- Storage: 500 GB Hard Disk
- Input Devices: Standard keyboard and mouse
- Display: SVGA monitor

Software Requirements

- Operating System: Windows 10 (minimum)
- Programming Language: Python 3.7+
- Key Libraries: scikit-learn 0.22+, LightGBM, pandas, NumPy, matplotlib, seaborn, tkinter
- Dataset: UCI Machine Learning Repository – Banknote Authentication (banknotes.csv)

VII. EXPERIMENTAL RESULTS AND EVALUATION

All experiments were conducted on a standard consumer-grade laptop running Windows 10, using only built-in CPU resources. The UCI Banknote Authentication dataset (1,372 records, 4 features, 1 class label) was used throughout. Three train-test split ratios were evaluated: 80:20, 70:30, and 60:40. Table II reports representative metrics at the 80:20 split, where differences between algorithms are most pronounced.

TABLE II. Algorithm Performance Comparison (80:20 Split)

Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
KNN	97.09	97.11	97.09	97.08
Naive Bayes	84.36	84.50	84.36	84.25
Decision Tree	98.18	98.19	98.18	98.18
SVM	99.27	99.28	99.27	99.27
Random Forest	99.27	99.30	99.27	99.27
Logistic Regression	97.82	97.83	97.82	97.81
LightGBM (Ext.)	**100.00**	**100.00**	**100.00**	**100.00**

A. Gesture Recognition Performance

KNN with k=2 achieved 97.09% accuracy under the 80:20 split. The algorithm performed consistently across Precision, Recall, and F1-Score, confirming that the four wavelet features form a well-separated feature space. However, KNN is sensitive to noise and its inference time scales with dataset size, making it less suitable for large-scale real-time deployments.

B. Voice Command Evaluation

Naive Bayes recorded the lowest accuracy at 84.36%, attributable to its independence assumption, which does not hold perfectly for the correlated wavelet features. While it provides a fast probabilistic baseline, its lower recall indicates a higher rate of undetected counterfeit notes, which is unacceptable in a production setting.

C. Multimodal Interaction Latency

Decision Tree achieved 98.18% accuracy under 80:20 and reached 100% under the 60:40 ratio. Its interpretable rule structure makes it attractive for regulatory environments where audit trails are required. It was retained as the default saved model for the test-data prediction module due to its balance of accuracy and interpretability.

D. Resource Utilization and Stability

Both SVM and Random Forest achieved 99.27% accuracy at the 80:20 split, with near-identical Precision, Recall, and F1-Scores. SVM's margin maximisation and RF's ensemble averaging both proved highly effective for this binary classification task. Random Forest exhibited lower variance across split ratios than SVM, suggesting better generalisation. Logistic Regression attained 97.82% accuracy, offering an interpretable probabilistic output.

E. User Experience Evaluation

LightGBM, deployed as the extension algorithm and trained on the full feature matrix, achieved 100% accuracy, Precision, Recall, and F1-Score under the 80:20 split. This result confirms that gradient-boosted trees with leaf-wise growth provide the strongest classification performance. Across all three split ratios, LightGBM sustained 100% accuracy, demonstrating superior generalisation over all competitors. In contrast, Naive Bayes showed the most variability (84–86%), and KNN and RF maintained consistently high accuracy (97–99%).

VIII. CONCLUSION

This paper presented a systematic framework for evaluating supervised machine learning algorithms applied to fake currency detection. Seven classifiers—KNN, Naive Bayes, Decision Tree, SVM, Random Forest, Logistic Regression, and the extended LightGBM—were benchmarked on the UCI Banknote Authentication dataset across three train-test split ratios. LightGBM achieved perfect classification accuracy of 100% under the 80:20 split, surpassing all traditional algorithms and demonstrating strong generalisation across all split conditions.

Naive Bayes was the weakest performer due to its independence assumption, while SVM and Random Forest both achieved 99.27% accuracy and represent strong alternatives where interpretability or variance stability is prioritised. Decision Tree offers a useful middle ground, reaching 100% accuracy under the 60:40 split with fully transparent decision rules.

Future work will explore additional advanced algorithms such as XGBoost and Multi-Layer Perceptrons (MLP), investigate feature engineering strategies to improve performance under noisy or degraded banknote scans, and extend the framework to multi-class denomination recognition and real-time video-based currency authentication pipelines.

IX. REFERENCES

[1] M. Aoba, T. Kikuchi, and Y. Takefuji, "Euro Banknote Recognition System Using a Three-layered Perceptron and RBF Networks," IPSJ Transactions on Mathematical Modeling and its Applications, May 2003.

[2] S. Desai, S. Kabade, A. Bakshi, A. Gunjal, M. Yeole, "Implementation of Multiple Kernel Support Vector Machine for Automatic Recognition and Classification of Counterfeit Notes," International Journal of Scientific & Engineering Research, October 2014.

[3] C. Gigliarano, S. Figini, P. Muliere, "Making classifier performance comparisons when ROC curves intersect," Computational Statistics and Data Analysis, vol. 77, pp. 300–312, 2014.

[4] E. Gillich and V. Lohweg, "Banknote Authentication," 2014.

[5] H. Hassanpour and E. Hallajian, "Using Hidden Markov Models for Feature Extraction in Paper Currency Recognition," 2010.

[6] Z. Huang, H. Chen, C. J. Hsu, W. H. Chen and S. Wu, "Credit rating analysis with support vector machines and neural network: a market comparative study," 2004.

[7] C. Kumar and A. K. Dudyala, "Banknote Authentication using Decision Tree rules and Machine Learning Techniques," ICACEA, 2015.

[8] M. Lee and T. Chang, "Comparison of Support Vector Machine and Back Propagation Neural Network in Evaluating the Enterprise Financial Distress," International Journal of Artificial Intelligence & Applications, vol. 1, no. 3, pp. 31–43, 2010.

[9] C. Nastoulis, A. Leros, and N. Bardis, "Banknote Recognition Based On Probabilistic Neural Network Models," WSEAS International Conference on SYSTEMS, Athens, Greece, July 2006.

[10] S. Omatu, M. Yoshioka and Y. Kosaka, "Bankcurrency Classification Using Neural Networks," IEEE, 2007.

[11] A. Patle and D. S. Chouhan, "SVM Kernel Functions for Classification," ICATE, 2013.

[12] E. L. Prime and D. H. Solomon, "Australia's plastic banknotes: fighting counterfeit currency," Angewandte Chemie International

edition, vol. 49, no. 22, pp. 3726–3736, May 2010.

[13] A. Roy, B. Halder, and U. Garain, "Authentication of currency notes through printing technique verification," ICVGIP, pp. 383–390, 2010.

[14] P. D. Shahare and R. N. Giri, "Comparative Analysis of Artificial Neural Network and Support Vector Machine Classification for Breast Cancer Detection," International Research Journal of Engineering and Technology, Dec 2015.

[15] F. Takeda, L. Sakoobunthu and H. Satou, "Thai Banknote Recognition Using Neural Network and Continuous Learning by DSP Unit," KES, 2003.

[16] M. Thirunavukkarasu et al., "Comparison of SVM and BPN methods in predicting protein virulence factors," Journal of Industrial Pollution Control, vol. 33, no. 2, pp. 11–19, 2017.

[17] C.-Y. Yeh, W.-P. Su, and S.-J. Lee, "Employing multiple-kernel support vector machines for counterfeit banknote recognition," Applied Soft Computing, vol. 11, no. 1, pp. 1439–1447, Jan. 2011.

[18] UCI Machine Learning Repository, Banknote Authentication Dataset. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/banknote+authentication

[19] V. K. Verma, A. Yadav, and T. Jain, "Key Feature Extraction and Machine Learning-Based Automatic Text Summarization," Emerging Technologies in Data Mining and Information Security, Springer, Singapore, pp. 871–877, 2019.