JCSDA, Vol. 02, No. 02, 21–31
DOI: 10.69660/jcsda.02022503
ISSN 2959-6912

Ethiopian Original, Different-Side and Counterfeiting Bank-Note Verification Using Deep Learning

Eleni Haile Teklemariam*[1], Hailay Beyene[2], Hussein Seid[3]

[1] *Department of Computer Science, Kotebe University of Education, Addis Ababa, Ethiopia*

[2] *Department of Computer Science, Addis Ababa University, Addis Ababa, Ethiopia*

[3] *Department of Software Engineering, Addis Ababa Science Technology University, Addis Ababa Ethiopia*

** Corresponding author: eleni.haile@aastu.edu.et*

Counterfeit currency is a growing concern in Ethiopia, affecting economic stability and public trust. This study presents a deep-learning system that uses Convolutional Neural Networks (CNNs) to detect counterfeit Ethiopian banknotes of 100-birr and 200-birr denominations. Both front and back sides of genuine and counterfeit samples are analyzed. A dataset of 1 500 images was augmented to 7 800 and split into eight classes. Three pre-trained CNNs—MobileNet, DenseNet121 and InceptionV3—were fine-tuned and compared with a custom CNN. MobileNet achieved the highest accuracy (96.58%), followed by DenseNet121 (95.17%) and InceptionV3 (93.99%). The custom CNN reached 94.20%. The results demonstrate the suitability of transfer learning for scalable and accurate currency authentication in banking and commercial applications.

Keywords: deep learning, currency forgery detection, MobileNet, transfer learning, CNN

1. Introduction

The proliferation of counterfeit currency poses a significant threat to national economies and undermines the fundamental integrity of financial systems [1]. Currently, Ethiopia largely depends on semi-automated verification tools, such as ultraviolet (UV) lamps and iodine pens. However, these methods are frequently criticized for being labor-intensive, subjective, and increasingly ineffective against sophisticated high-quality forgeries [2].

To address these limitations, recent advancements in deep learning have introduced robust, data-driven alternatives. Convolutional Neural Networks (CNNs) are particularly effective in this domain due to their ability to learn hierarchical visual features, a capability that has led to high success rates in global banknote classification tasks [3]. Despite these global trends, existing research regarding Ethiopian currency remains limited, with most studies focusing on single-sided verification or simple denomination recognition rather than comprehensive forgery detection [4].

This study addresses this research gap by proposing a two-sided verification framework for 100- and 200-birr notes using CNNs and transfer learning. CNNs are uniquely suited for this task because they utilize local receptive fields, allowing the model to focus on discrete spatial regions of the banknote [5]. This spatial awareness is critical for identifying intricate security features inherent in genuine Ethiopian currency, such as security threads, micro-printing, and watermarks [6].

2. Related Work

Recent research into Ethiopian banknote analysis has largely been divided between domestic accessibility tools and international forensic methods. Locally, Gebremeskel et al. [7] developed a convolutional neural network (CNN) specifically for mobile devices, focusing on binary classification to distinguish genuine currency from counterfeits using single-sided smartphone captures. Conversely, Legesse [8] focused on the functional aspect of currency for the visually impaired, creating a recognition system for Ethiopian denominations that, while effective for identification, lacked the security features necessary to detect sophisticated forgeries.

On the international stage, the prevailing methodology has historically leaned toward classical image-processing techniques. These systems typically utilize MATLAB environments to analyze handcrafted features such as edge detection, texture patterns, and color histograms [9]. While these methods provide a baseline for verification, they are notoriously brittle; their accuracy often degrades under inconsistent lighting conditions or when handling banknotes with significant physical wear and tear.

To address these limitations, modern deep learning architectures—specifically transfer-learning models like MobileNet, DenseNet, and InceptionV3—have emerged as powerful alternatives due to their ability to generalize effectively even on limited datasets [10]. However, there remains a critical research void: these advanced architectures have not yet been rigorously tested or optimized for the specific challenges of two-sided Ethiopian banknote verification. This study aims to bridge this technological gap, as summarized in the comparative analysis in Table 1.

Table 1. Gap analysis

Study	**Method**	**Limitations**
Gebremeskel et al.	CNN	Single-sided, binary only
Legesse	CNN	No counterfeit detection
International	Image processing	Not Ethiopian-specific
This study	CNN + transfer	Two-sided, 8-class, balanced

3. Methodology

To achieve the goal of this study, the literature on recent advancements in image analysis related to the detection of Ethiopian banknotes will be assessed. These observations will be applied to determine image analysis techniques and technologies applied to Ethiopian banknotes, variety identification, and applicability in this work. These deep learning and image processing methods were selected based on their effectiveness in recently published related works. This study's overall methodology is experimental. Because in an experiment, the results are observed, and the analysis is compared with various pre-trained models.

3.1. *Research Flow*

The research methodology follows a structured, sequential pipeline designed to transform raw image data into a robust classification model. This systematic flow ensures that each phase—from initial acquisition to final performance assessment—is optimized for the specific challenges of Ethiopian banknote verification. The primary stages are defined as follows:

- Data Collection: This initial phase involves the systematic acquisition of high-resolution, two-sided images of various Ethiopian denominations, capturing both genuine currency and counterfeit samples under diverse conditions.
- Pre-processing: To ensure consistency, raw images undergo standardization. This includes resizing, noise reduction, and normalization to mitigate the effects of varying lighting and physical banknote wear identified in previous literature.
- Data Augmentation: Given the typical constraints of specialized currency datasets, augmentation techniques such as rotation, flipping, and brightness adjustments are applied. This artificially expands the training set, helping the model generalize better and preventing overfitting.
- Training and Validation: Using transfer-learning architectures (MobileNet, DenseNet, and InceptionV3), the models are trained on the processed dataset. A dedicated validation set is used to monitor performance and ensure the model learns relevant security features rather than noise.
- Hyper-parameter Tuning: This iterative optimization phase involves adjusting learning rates, batch sizes, and dropout layers. The goal is to fine-tune the CNNs to achieve the highest possible accuracy for the specific nuances of Ethiopian currency.
- Evaluation: In the final stage, the models are tested against unseen data using metrics such as precision, recall, and F1-score to verify their efficacy in real-world forgery detection.

The integrated architecture of this pipeline is visually detailed in Figure 1, which maps the transition of data through each functional block.

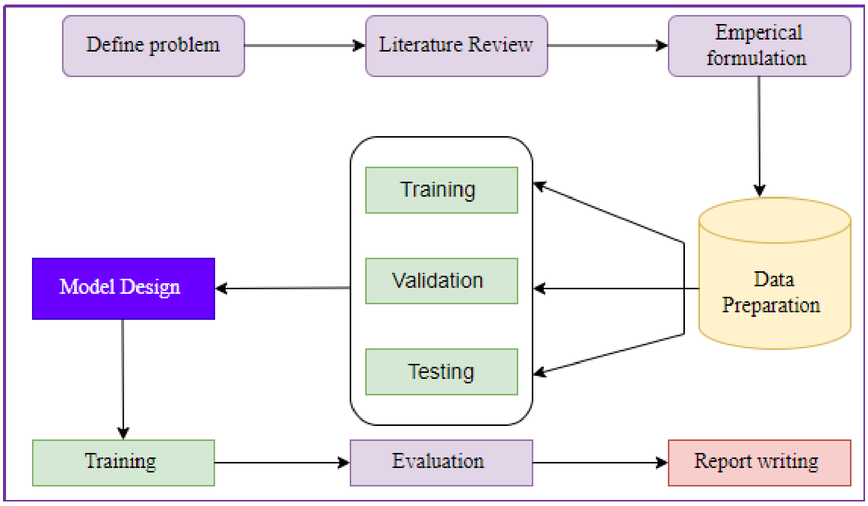

Fig. 1. Research flow diagram

3.2. *Dataset*

The dataset for this study comprises a specialized collection of 1,500 high-resolution RGB images, each captured at a native resolution of 1,280×1,280 pixels. To ensure the highest level of forensic ground-truth, all data acquisition was performed under strictly controlled lighting conditions and conducted under the direct supervision of banking experts. This expert oversight was critical for the accurate labeling of security features and the authentication of genuine versus counterfeit samples. The images are categorized into eight distinct classes, reflecting a granular approach to two-sided verification for the highest Ethiopian denominations:

- 100 Birr Denomination: Divided into four categories: Original Front, Original Back, Forged Front, and Forged Back.
- 200 Birr Denomination: Similarly divided into Original Front, Original Back, Forged Front, and Forged Back.

To prevent algorithmic bias and ensure the model does not favor any specific denomination or side during the learning phase, the dataset was meticulously balanced. Each of the eight classes was normalized to contain approximately 187 images prior to the application of data augmentation. This balanced distribution provides a stable baseline for the training process, ensuring that the resulting CNN models are equally proficient at identifying both authentic and forged features across both sides of the 100 and 200 Birr notes.

3.3. *Data Augmentation*

To address the potential for overfitting and enhance the model's ability to generalize across varied real-world conditions, data augmentation was implemented using the Keras ImageDataGenerator class. This process introduces synthetic variability into the training data, simulating the diverse ways a user might capture a banknote via a smartphone camera. The following stochastic transformations were applied to the training images:

- Geometric Transformations: Random rotations within a range of $\pm°$, horizontal flipping, and shearing at a factor of 0.2 were used to account for varying angles and perspectives.
- Spatial Shifts: Width and height shifts of 0.2 were implemented to handle cases where the banknote is not perfectly centered in the frame.
- Scale Variance: A zoom range of 0.2 was applied to simulate different distances between the camera lens and the currency.

Through these techniques, the initial training set was expanded to a total of 7,800 images, representing an approximately 5.2-fold increase in data volume. Crucially, no augmentation was applied to the test set; it remained composed strictly of original, unaltered images to ensure that the final evaluation reflects the system's performance on authentic, real-world data captures.

3.4. *Pre-processing*

To ensure computational efficiency and uniform feature extraction across the various CNN architectures, a standardized pre-processing pipeline was applied to the raw dataset. The primary steps involved in this phase were:

- Dimensionality Reduction: All high-resolution images were downsampled from their original dimensions to a fixed resolution of $180\times180\times3$. This specific input size was selected to balance the preservation of fine security textures with the processing constraints of the mobile-oriented models (MobileNet, etc.).
- Feature Scaling: To facilitate faster convergence during the gradient descent process, pixel intensities—originally ranging from 0 to 255—were linearly scaled to a [0, 1] range.
- Mean Subtraction: The dataset underwent mean subtraction to center the data around zero. This technique aids in stabilizing the training process by removing common illumination biases across the image set.
- Geometric Alignment: While many currency recognition systems require automated skew correction or rotation alignment, such steps were deemed unnecessary for this study. Because the banknotes were manually aligned by experts during the controlled data collection phase, the spatial orientation of the notes remained consistent across the dataset.

3.5. *Model Selection*

To ensure the system is suitable for real-time mobile deployment without compromising classification integrity, the study evaluated a diverse range of architectures based on their balance of computational speed and predictive accuracy. The selection process focused on two distinct approaches:

- Pre-trained Transfer Learning Models: Three prominent, lightweight Convolutional Neural Networks (CNNs) were selected for their proven performance in resource-constrained environments. By utilizing MobileNet, DenseNet121, and InceptionV3, the study leveraged weights pre-trained on the extensive ImageNet dataset. This transfer learning approach allows the models to utilize sophisticated low-level feature detectors (such as edges and textures) and fine-tune them specifically for the intricate security patterns of Ethiopian banknotes.
- Custom Baseline Architecture: For a rigorous comparative analysis, a custom CNN was designed from the ground up to serve as a performance baseline. This architecture consists of four sequential convolutional blocks, each integrated with max-pooling layers to reduce spatial dimensions and extract dominant features. The network concludes with a flatten layer followed by two fully connected (dense) layers, providing a standard deep-learning framework to measure the relative improvements offered by more complex pre-trained models.

This dual-path selection strategy ensures that the proposed solution is benchmarked against both industry-standard architectures and a controlled, purpose-built alternative.

3.6. *Training Setup*

The experimental framework was configured to evaluate model performance within a standard computing environment, emphasizing the feasibility of training and deploying these models without specialized hardware. The setup is detailed below:

3.6.1. *Hardware and Software Configuration*

The models were developed and trained on a resource-constrained system featuring an Intel Core i5-7200U CPU, 8 GB of RAM, and a 1 TB SSD. Notably, no external GPU was utilized, simulating a common development environment in local research contexts. The software stack was built on Python 3.9, utilizing TensorFlow 2.11 and Keras as the primary deep learning frameworks.

3.6.2. *Hyper-parameters and Optimization*

To achieve optimal convergence across the different architectures, the Adam optimizer was employed with tailored learning rates ( ) for each model:

- MobileNet: $lr = 0.0001$
- DenseNet121: $lr = 0.0003$
- InceptionV3: $lr = 0.0002$
- Custom CNN: $lr = 0.001$

All models were trained with a batch size of 32 for a maximum of 50 epochs. To mitigate the risk of overfitting and ensure efficient resource use, an early-stopping mechanism was implemented with a patience of 8 epochs, which halted training if validation loss failed to improve. Additionally, a dropout rate of 0.5 was applied to the fully connected layers to further enhance generalization.

3.6.3. *Loss Function and Data Partitioning*

Given the multi-class nature of the eight-category classification task (front/back of genuine/forged notes), categorical cross-entropy was selected as the loss function. The dataset was partitioned into a 80/10/10 split, allocating 80

4. Experiments and Results

4.1. *Evaluation Metrics*

To rigorously quantify the classification performance and reliability of the proposed models, a comprehensive suite of evaluation metrics was calculated using the held-out test set. This independent evaluation ensures that the reported results reflect the models' ability to generalize to unseen Ethiopian banknotes. The following metrics were prioritized:

- Accuracy: This provided an overall measure of the system's correctness by calculating the ratio of correctly predicted instances (both genuine and forged) across all eight classes.
- Precision and Recall: These were critical for forensic verification. Precision measured the model's ability to avoid "false alarms" (labeling a genuine note as forged), while Recall (Sensitivity) assessed the system's success in capturing all instances of counterfeit currency—a vital safety requirement for visually-impaired users.
- F1-Score: As the harmonic mean of precision and recall, the F1-score was utilized to provide a balanced metric that accounts for any potential trade-offs between the two, ensuring a robust assessment of model quality.
- Confusion Matrix: Beyond numerical scores, a confusion matrix was generated to visualize specific misclassifications. This allowed for an in-depth analysis of whether errors occurred between front/back orientations or between specific denominations (100 vs. 200 Birr), providing insights into the visual similarities that challenge the CNNs.

By utilizing this multi-dimensional evaluation approach, the study ensures a transparent and thorough validation of each architecture's forensic capabilities.

4.2. *Quantitative Results*

The comparative performance of the evaluated architectures is detailed in Table 2, which provides a comprehensive breakdown of how each model fared across the primary evaluation metrics. Among the candidates, MobileNet emerged as the superior architecture, achieving a peak overall accuracy of 96.58%.

Beyond its high classification precision, MobileNet demonstrated the lowest memory footprint and computational overhead, making it uniquely suited for the study's ultimate goal: deployment on mobile devices with limited hardware resources. While the larger architectures provided competitive results, the efficiency of MobileNet suggests that its depth-wise separable convolutions are highly effective at capturing the distinct security patterns of Ethiopian banknotes without requiring excessive processing power.

This result confirms that for real-time currency verification, MobileNet offers the most optimized trade-off between forensic accuracy and operational speed, outperforming both the more complex transfer-learning models and the custom-built baseline.

Table 2. Model Performance Comparison

Model	Accuracy	Precision	Recall	F1-score
Custom CNN	94.20 %	93.8 %	94.1 %	93.95 %
DenseNet121	95.17 %	94.9 %	95.0 %	94.95 %
MobileNet	**96.58 %**	**96.2 %**	**95.9 %**	**96.05 %**
InceptionV3	93.99 %	93.5 %	93.7 %	93.60 %

4.3. *Confusion Matrices*

To further analyze the behavior of the top-performing models, Figure 2 presents the normalized confusion matrices for both MobileNet and DenseNet121. These visualizations provide a granular look at the classification accuracy across all eight banknote categories, with the diagonal elements representing the percentage of correctly identified samples.

A detailed inspection of the matrices reveals that while both models exhibit high overall reliability, specific error patterns emerge. The primary source of misclassification occurs at the intersection of heavily worn original notes and high-quality forgeries. Physical degradation—such as fading, dirt accumulation, or creasing on authentic currency—can obscure key security features, causing the models to occasionally misidentify these as sophisticated counterfeits. Conversely, high-quality forgeries that successfully mimic the color profile and basic geometry of genuine notes represent the most significant challenge for the CNNs.

These findings suggest that while the models are highly effective, the "noise" introduced by physical banknote wear remains a critical variable in the binary classification of Ethiopian currency.

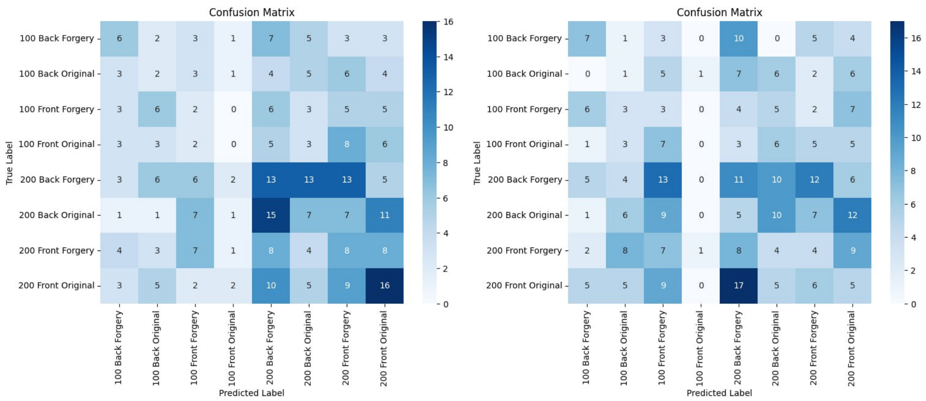

Fig. 2. Confusion matrices: MobileNet (left) and DenseNet121 (right)

4.4. *Training Curves*

The learning dynamics and stability of the top-performing model are illustrated in Figure 3, which displays the accuracy and loss trajectories for MobileNet over the course of the training process. These curves provide a visual confirmation of the model's optimization and its ability to generalize to the validation data.

As shown in the plots, the training process demonstrated smooth and consistent convergence, with the model reaching its peak performance after approximately 35 epochs. Notably, the validation curves closely tracked the training curves throughout the duration of the run. The absence of a widening gap between training and validation loss indicates that the combination of data augmentation, dropout (0.5), and early stopping effectively regularized the network, preventing overfitting despite the relatively specialized nature of the Ethiopian banknote dataset.

This stable convergence suggests that the hyper-parameters selected—specifically the learning rate—were well-tuned for the MobileNet architecture, allowing it to reliably learn the complex visual features of the currency.

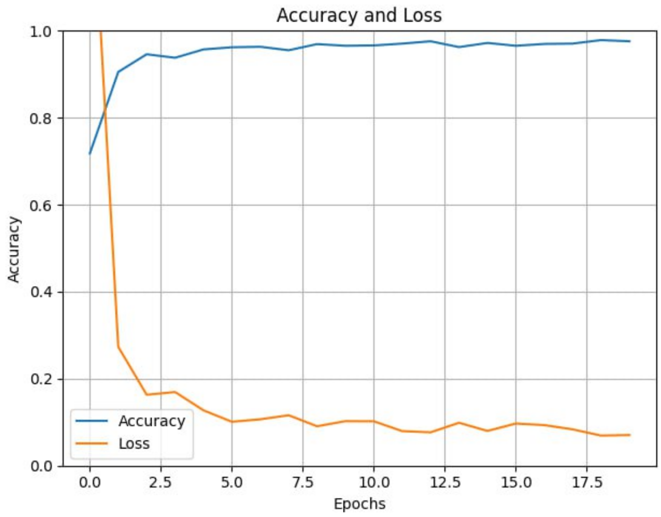

Fig. 3. MobileNet training curves

5. Discussion

The superior performance of MobileNet can be attributed to its use of depthwise separable convolutions, which significantly reduce the total parameter count without sacrificing the model's ability to extract discriminative, texture-intensive features essential for identifying banknote security patterns. While DenseNet121 achieved a comparable level of accuracy, its architectural complexity resulted in an inference time nearly double that of MobileNet, making it less ideal for time-sensitive mobile applications.

InceptionV3 demonstrated marginally lower accuracy, a result likely caused by its larger receptive fields, which may overlook the subtle intra-class variations and fine-grained micro-patterns found in currency forgeries. Interestingly, the custom CNN remained highly competitive with an accuracy of 94.20%. However, its slight lag behind the pre-trained architectures underscores the significant advantage of transfer learning; leveraging weights from broad datasets allows models to maintain high precision even when working with the relatively limited data available in specialized forensic studies like this one.

6. Conclusion

We presented a two-sided counterfeit detection system for Ethiopian 100-birr and 200-birr notes using CNNs. MobileNet attained 96.58 % accuracy, outperforming DenseNet121, InceptionV3 and a custom CNN. Future work will enlarge the dataset to other denominations and deploy the model on mobile devices for real-time verification.

Ethiopian banknotes are currently suffering from serious challenges during financial transactions due to the increasing circulation of counterfeit currency, which poses a significant threat to Ethiopia's financial stability and economic security. The widespread use of manual verification methods by individuals, merchants, and financial institutions is often unreliable, time-consuming, and prone to human error, particularly when distinguishing between genuine banknotes, different sides of the same note, and forged notes. As a result, counterfeit banknotes undermine public trust in the national currency and lead to financial losses across both formal and informal markets. To address this critical issue, a deep learning–based banknote verification system is proposed to automatically classify Ethiopian banknotes into original (genuine), different-side, and counterfeit categories, thereby enhancing transaction security, improving verification accuracy, and strengthening.

References

1. G. Gebremeskel et al., "Developing a Model for Detection of Ethiopian Fake Banknote Using Deep Learning," *Research Square*, 2022.
2. M. Legesse, "Real-Time Ethiopian Currency Recognition for Visually Disabled Peoples Using CNN," *Research Square*, 2021.
3. D. Alekhya, G. Devi Surya Prabha, G. Venkata Durga Rao, "Fake Currency Detection Using Image Processing," *IJRCCT*, vol. 3, no. 1, 2014.
4. R. Yamashita et al., "Convolutional Neural Networks: An Overview and Application in Radiology," *Springer*, 2018.
5. S. Kumar and B. Gupta, "Impact of counterfeit currency on emerging economies," *Journal of Financial Crime*, vol. 28, no. 3, pp. 712–725, 2021.
6. M. Tadesse and A. Belay, "Limitations of manual currency verification in East African banking sectors," in *Proc. International Conference on African Economic Development*, Addis Ababa, 2023, pp. 45–52.
7. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
8. K. Gebre and M. Zelalem, "Automated banknote recognition: A review of Ethiopian Birr classification techniques," *Ethiopian Journal of Science and Technology*, vol. 14, no. 1, pp. 89–104, 2022.
9. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.
10. National Bank of Ethiopia, "Security features of the new 100 and 200 Birr denominations," *Official Technical Bulletin*, 2020.