iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

# COUNTERFEIT PRODUCT DETECTOR USING DEEP LEARNING

**S. Divakar**
Research Scholar, School of Computing Sciences, VISTAS, Chennai

**Dr. Ragahvendran**
Professor, School of Computing Sciences, VISTAS, Chennai

**ABSTRACT:**
Counterfeit products have become a major issue in global markets, causing financial losses for companies and safety risks for consumers. Detecting fake products manually is difficult, time-consuming, and often inaccurate. To address this problem, this project proposes a Counterfeit Product Detector using Deep Learning, which automatically identifies whether a product is genuine or counterfeit using image analysis. The proposed system uses deep learning models, particularly Convolutional Neural Networks (CNN), to analyze product images and detect subtle differences between original and fake items. The model is trained using a dataset containing images of both authentic and counterfeit products. By learning visual features such as logos, packaging patterns, colors, and textures, the system can accurately classify products. The system consists of several modules including image input, preprocessing, feature extraction, deep learning classification, and result prediction. When a user uploads a product image, the system processes the image and predicts whether the product is original or counterfeit.

**Keywords**
Counterfeit Detection, Deep Learning, Convolutional Neural Network (CNN), Image Classification, Product Authentication, Transfer Learning, Visual Feature Extraction, Fake Product Detection, Computer Vision.

## INTRODUCTION

Counterfeit goods represent one of the most pervasive and economically damaging phenomena in global commerce today. From luxury fashion items and pharmaceutical drugs to electronic components and everyday consumer products, counterfeit merchandise infiltrates virtually every product category. The consequences extend well beyond financial loss for legitimate manufacturers. Counterfeit pharmaceuticals can endanger patient health and lives. Fake electronic components can cause device failures with serious safety implications. Substandard food and beverage counterfeits expose consumers to health risks that may be invisible until harm has already occurred. According to estimates from international trade organizations, the global trade in counterfeit and pirated goods has grown into a multi-trillion-dollar problem that undermines supply chain integrity, erodes consumer trust, and deprives governments of tax revenue.

Traditional methods of counterfeit detection depend heavily on the expertise of trained inspectors who examine products for telltale visual signs of inauthenticity such as misaligned logos, incorrect typography, substandard printing quality, abnormal material texture, and deviations from authentic packaging specifications. While such expert inspection can be effective, it is inherently slow, expensive, inconsistent across individuals, and entirely impractical at the volumes of product units that flow through modern supply chains and e-commerce platforms every day. Counterfeiters have also become increasingly sophisticated, reproducing security features such as holograms, QR codes, and embossed seals with high fidelity, further reducing the reliability of manual inspection methods.

The rapid advancement of deep learning and computer vision provides a compelling and timely solution to this problem. Convolutional Neural Networks have demonstrated extraordinary capability in learning discriminative visual representations directly from raw image data, capturing subtle differences in texture, color, spatial pattern, and structural geometry that human inspectors may be unable to perceive or describe. By training a CNN on a carefully curated dataset of authentic and counterfeit product images across multiple product categories, it becomes possible to construct an automated classification system that accepts a product photograph as input and returns a prediction of genuine or counterfeit with a quantified confidence level.

This paper presents the Counterfeit Product Detector using Deep Learning, a comprehensive end-to-end system that integrates transfer learning from pretrained backbone networks, multi-module image preprocessing, and a user-accessible web interface for real-time product authentication. The system is designed to be product-category

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

agnostic and retrainable for any product domain by supplying domain-specific training images. The paper is organized as follows. Section II reviews related work in deep learning-based image classification and product authentication. Section III describes the limitations of existing systems. Section IV presents the system architecture. Section V explains the proposed methodology. Section VI describes the system modules. Section VII discusses implementation. Section VIII presents experimental evaluation results. The final section presents conclusions and future directions.

## RELATED WORK

Research on automated product authentication and counterfeit detection has evolved through several distinct phases. Early work relied on handcrafted feature descriptors such as Scale-Invariant Feature Transform and Histogram of Oriented Gradients combined with classical machine learning classifiers including Support Vector Machines and k-Nearest Neighbour algorithms. These approaches demonstrated the viability of image-based authentication in controlled laboratory settings but exhibited significant sensitivity to variations in lighting, viewing angle, and image resolution that are unavoidable in practical deployment environments.

The transformative impact of deep learning became evident following the introduction of AlexNet by Krizhevsky, Sutskever, and Hinton in 2012, which demonstrated that deep convolutional architectures trained on large labelled datasets could learn hierarchical visual representations far more powerful and generalizable than any manually designed feature extractor. Subsequent CNN architectures including VGGNet, GoogLeNet, and ResNet progressively improved classification accuracy and established very deep residual networks as the dominant paradigm for visual recognition tasks. These architectures, when pretrained on the ImageNet dataset comprising over one million images across one thousand categories, provide powerful general-purpose feature extractors that can be fine-tuned for domain-specific tasks using comparatively small amounts of labelled training data through transfer learning.

Transfer learning has been particularly valuable in the product authentication domain because large-scale counterfeit product datasets are difficult and expensive to collect and label. Wilber et al. demonstrated that ResNet features pretrained on ImageNet and fine-tuned on domain-specific brand imagery substantially outperformed both handcrafted descriptors and networks trained from scratch on small domain datasets for logo recognition tasks. Gao et al. proposed a Siamese network architecture for product authentication that learns an embedding space in which authentic and counterfeit product images are maximally separated, enabling effective matching even for product categories with limited training examples.

More recent research has focused on improving both the accuracy and interpretability of deep learning-based authentication. Selvaraju et al. introduced Gradient-weighted Class Activation Mapping, which generates visual heatmaps indicating which regions of an input image most strongly influenced the model's classification decision. This interpretability mechanism is particularly important in authentication contexts where decisions may carry legal or regulatory significance and must be explainable to non-technical stakeholders. Tan and Le introduced EfficientNet, which applies compound scaling of network depth, width, and input resolution, achieving higher accuracy per parameter than conventional scaling strategies and making it suitable for deployment in resource-constrained environments. These developments directly inform the design of the proposed Counterfeit Product Detector.

Research specifically addressing pharmaceutical tablet authentication using CNN features has demonstrated that subtle differences in surface texture, colour uniformity, coating thickness, and edge sharpness between authentic and counterfeit tablets, which are invisible to the naked eye, are reliably captured by intermediate convolutional layer activations. Huang et al. proposed a multi-scale attention CNN for product authentication that weights feature map activations according to their informational contribution, providing interpretable attention maps aligned with physically meaningful product authentication criteria such as logo placement precision and packaging colour fidelity. These approaches confirm that deep learning is not merely a black-box classifier but can be guided to attend to the same visual cues that trained human experts examine.

## EXISTING SYSTEM

In many organizations and retail environments, counterfeit product detection is still performed through manual inspection by trained personnel or relies on authentication technologies embedded in product packaging. Manual expert inspection involves trained staff examining products for visual indicators of counterfeiting such as logo misalignment, incorrect font weight, substandard print quality, abnormal surface texture, and packaging dimension deviations. While experienced inspectors can achieve reasonable accuracy for familiar product categories, this approach is fundamentally limited by human perceptual capacity, cannot scale to the volume of

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

units processed in large retail and e-commerce operations, and produces inconsistent results across different inspectors and inspection sessions.

Serialization and track-and-trace systems represent another widely deployed approach, assigning unique identifiers encoded in barcodes, QR codes, or RFID tags to individual product units at the point of manufacture. These systems enable authenticity verification by querying a central database at any point in the supply chain. However, they are vulnerable to identifier cloning, where counterfeiters copy valid identifiers from authentic products, and to supply chain fraud involving the acquisition of unused identifier ranges. They also provide limited protection at the consumer level, as most consumers lack the technical means to perform database verification at the point of purchase.

Physical authentication features such as holograms, colour-shifting inks, microprinting, and embedded security threads are widely incorporated into the packaging of high-value product categories. These features are designed to be economically prohibitive for counterfeiters to reproduce faithfully. However, advances in printing technology and material science continue to reduce the cost of replicating even sophisticated security features, and the ability of ordinary consumers to assess the authenticity of such features without specialist training or equipment is severely limited.

The existing system therefore suffers from several fundamental limitations that the proposed deep learning solution is designed to address. These limitations include:

- Dependency on manual inspection that is slow, subjective, and inconsistent
- Inability to scale to the volumes of products in modern supply chains and e-commerce
- Vulnerability of serialization systems to identifier cloning and supply chain fraud
- Progressive erosion of physical security feature effectiveness as replication technology improves
- Absence of an accessible, automated authentication mechanism for end consumers

These limitations collectively motivate the development of a deep learning-based automated counterfeit detection system that can process product images rapidly, consistently, and at scale while providing quantified confidence estimates and interpretable visual explanations for each authentication decision.

Figure 1: Existing System Flow — Physical Product → Manual Expert Inspection → Accept or Reject (slow, subjective, unscalable, no consumer access)

## SYSTEM ARCHITECTURE

The proposed Counterfeit Product Detector follows a four-layer pipeline architecture in which each layer performs a specific processing responsibility while passing structured outputs to the next layer. This layered organization ensures modularity, maintainability, and independent testability of each component. The four layers are the Input and Validation Layer, the Preprocessing and Feature Extraction Layer, the Classification and Explanation Layer, and the Result Presentation and Feedback Layer.

The Input and Validation Layer is the entry point of the system and is responsible for accepting product images submitted through the web interface or REST API. This layer validates the format, minimum resolution, and file size of each submission, rejects corrupted or unsupported files with informative error messages, and queues validated images for asynchronous processing. By enforcing input standards at this layer, the downstream processing pipeline is protected from malformed inputs that could cause processing errors or degraded classification accuracy.

The Preprocessing and Feature Extraction Layer prepares validated input images for classification by applying a standardized transformation pipeline. Each image is centre-cropped, resized to 224 by 224 pixels using bilinear interpolation, and normalized per channel using the ImageNet mean and standard deviation values. The pretrained ResNet-50 backbone processes the normalized image tensor through its convolutional and pooling layers, producing a 2048-dimensional feature vector that encodes a hierarchical representation of the image content. A parallel EfficientNet-B4 pathway processes the same image at 380 by 380 pixel resolution, producing a complementary feature representation that captures finer textural detail.

The Classification and Explanation Layer maps the extracted feature vectors to class probability scores using fully connected classification heads trained on the counterfeit product dataset. An ensemble averaging operation combines the softmax outputs of the two backbone pathways to produce the final calibrated probability estimate. Gradient-weighted Class Activation Mapping is applied to the final convolutional layer of each backbone to generate a visual attention heatmap that is upsampled to the original image resolution and overlaid on the input image, highlighting the regions most influential to the classification decision.

The Result Presentation and Feedback Layer delivers the classification label, calibrated confidence score, and Grad-CAM heatmap to the user through the web interface. Users can view the heatmap overlay to understand the

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

visual basis of the classification decision. An expert feedback mechanism allows authorized users to submit corrective labels for ambiguous or erroneous classifications, which are stored in a feedback database and used periodically to fine-tune the model through the continuous learning pipeline.

Figure 2: System Architecture — Input Layer → Preprocessing Layer → ResNet-50 + EfficientNet-B4 → Grad-CAM + Ensemble Output → Feedback Layer

## PROPOSED METHODOLOGY

The methodology of the Counterfeit Product Detector is organized around the principle of automated visual authentication through deep feature learning, ensemble classification, and interpretable explanation generation. The first stage of the methodology involves dataset construction and curation. Authentic and counterfeit product images are collected across five representative product categories: luxury handbags, pharmaceutical tablets, consumer electronics accessories, branded footwear, and packaged food products. Each image is reviewed and labelled by domain experts before inclusion in the training dataset, ensuring high label quality. Synthetic augmentation using geometric transformations, colour jitter, and compression simulation is applied to expand the training distribution and improve model robustness to real-world imaging variability.

The second stage concerns transfer learning and backbone fine-tuning. ResNet-50 and EfficientNet-B4 networks pretrained on the ImageNet dataset are used as starting points, with their final fully connected layers replaced by two-unit classification heads. The backbone weights are unfrozen and fine-tuned on the counterfeit product dataset using layer-wise learning rate decay, applying smaller learning rates to earlier convolutional layers that encode low-level general features and larger learning rates to later layers that encode domain-specific discriminative features. This strategy preserves the general visual representations learned during ImageNet pretraining while specializing the higher-level features for product authentication.

The third stage involves ensemble construction and confidence calibration. The softmax output probabilities of the two fine-tuned backbone models are averaged to produce the ensemble prediction. Temperature scaling is applied to the pre-softmax logits of each model individually to calibrate the probability outputs, reducing the systematic overconfidence that deep neural networks commonly exhibit. The temperature scaling parameter for each model is optimized on a held-out validation set to minimize the Expected Calibration Error, ensuring that the reported confidence scores accurately reflect the empirical likelihood of correct classification.

A two-tier decision protocol governs the final output of the system. Predictions with an authentic probability above 0.75 or below 0.25 are returned as automated high-confidence classifications. Predictions with an authentic probability between 0.40 and 0.60, indicating genuine uncertainty, are routed to an expert review queue where a human specialist examines the input image alongside the Grad-CAM heatmap to make a final determination. This hybrid protocol ensures that automated outputs are reliable while reserving human expertise for the genuinely ambiguous cases where automated systems are most prone to error.

Figure 3: Methodology Flow — Dataset Collection → Preprocessing → Transfer Learning Fine-Tuning → Ensemble + Calibration → Two-Tier Decision Output

The continuous learning methodology completes the lifecycle by incorporating expert-corrected classifications back into the training pipeline. Corrective labels accumulated in the feedback database are combined with a stratified sample of the original training data and used to fine-tune the model nightly at a reduced learning rate, preventing catastrophic forgetting of previously learned representations while improving accuracy on difficult or novel counterfeit variants. An automated accuracy gate prevents deployment of any updated model that falls below the minimum accuracy threshold on the original test set.

## MODULES DESCRIPTION

The system is organized into six functional modules that together form the complete counterfeit detection pipeline. Each module performs a specific processing responsibility and communicates with adjacent modules through well-defined data interfaces, enabling independent testing and replacement without disrupting the overall system.

**Image Input and Validation Module**

The Image Input and Validation Module serves as the system gateway and is responsible for accepting product images submitted through the web interface or REST API endpoint. It validates each submission for format compatibility, checking that the file is a JPEG, PNG, or WebP image. It enforces a minimum resolution of 224 by 224 pixels and a maximum file size of 10 megabytes. Corrupted files that fail checksum verification are rejected with descriptive error messages. Validated images are assigned a unique task identifier and placed in the processing queue for asynchronous handling by the preprocessing module. This separation of input handling from processing ensures that the web interface remains responsive under concurrent request load.

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

**Preprocessing and Normalization Module**

The Preprocessing and Normalization Module applies a standardized image transformation pipeline to bring all input images into the distribution expected by the pretrained backbone networks. Each validated image is first converted to the RGB colour space to discard any alpha channel information. A centre-crop operation removes potentially uninformative border regions before the image is resized to 224 by 224 pixels using bilinear interpolation for the ResNet-50 pathway and 380 by 380 pixels for the EfficientNet-B4 pathway. Per-channel normalization using the ImageNet mean and standard deviation values aligns the input statistics with those of the pretraining data, maximizing the utility of pretrained weight initializations. At inference time, test-time augmentation generates ten augmented variants of each image through horizontal flipping, rotation, and brightness jitter, and the final class probabilities are averaged across all ten variants to reduce prediction variance.

**Feature Extraction Module (CNN Backbones)**

The Feature Extraction Module implements the primary deep learning inference pipeline using two complementary CNN backbone architectures. The ResNet-50 backbone processes 224 by 224 pixel normalized image tensors through 50 convolutional and pooling layers organized as residual blocks that enable gradient flow through skip connections during backpropagation. The network's global average pooling layer produces a 2048-dimensional feature vector encoding hierarchical visual representations from low-level edge and colour features in early layers to high-level semantic object-part patterns in later layers. The EfficientNet-B4 backbone processes images at 380 by 380 pixel resolution and applies compound scaling of network depth, width, and resolution, producing complementary fine-grained textural features that enhance ensemble diversity. Both models run on GPU using PyTorch with CUDA acceleration, achieving an average combined inference throughput of approximately 22 images per second.

**Ensemble Classification and Calibration Module**

The Ensemble Classification and Calibration Module combines the per-class softmax outputs of the ResNet-50 and EfficientNet-B4 backbones by element-wise averaging to produce a single ensemble probability distribution over the authentic and counterfeit classes. Prior to averaging, each model's logit outputs are divided by a learned temperature scaling parameter optimized on the validation set to minimize Expected Calibration Error. The final ensemble probability for the counterfeit class serves as the classification confidence score reported to the user. Predictions are routed to the automated output path if the confidence exceeds the high-confidence threshold, or to the expert review queue if the confidence falls within the uncertainty band, implementing the two-tier decision protocol.

**Grad-CAM Visual Explanation Module**

The Grad-CAM Visual Explanation Module generates a visual attention heatmap for each classified image by computing the gradient of the target class score with respect to the final convolutional layer activations of each backbone network. The gradient magnitudes are used to weight the activation map channels, and a ReLU operation is applied to the weighted sum to retain only positive contributions to the classification decision. The resulting heatmap is upsampled to the original input image resolution using bilinear interpolation and overlaid as a colour gradient ranging from blue for low influence to red for high influence regions. The three rectangular bounding boxes enclosing the highest-activation regions are extracted and returned alongside the heatmap as explicit indicators of the image regions most informative for the authentication decision, assisting expert reviewers in focusing their attention efficiently.

**Continuous Learning and Feedback Module**

The Continuous Learning and Feedback Module implements a feedback loop in which expert-corrected classification labels are used to periodically improve the model's accuracy on difficult or novel counterfeit variants. Expert corrections submitted through the web interface are logged in the feedback database with the original image reference, the model's original prediction, the expert's corrected label, and the submission timestamp. Nightly, a retraining pipeline retrieves all new corrective labels accumulated since the previous cycle, combines them with a stratified twenty percent sample of the original training dataset, and fine-tunes the ensemble models for five epochs using a reduced learning rate of one times ten to the negative five to prevent catastrophic forgetting. Updated models are evaluated against the full original test set, and deployment proceeds only if accuracy exceeds the minimum threshold of ninety-five percent.

## IMPLEMENTATION

The Counterfeit Product Detector is implemented in Python 3.10 and deployed as a Docker containerized web application using a multi-stage build process. The core deep learning inference pipeline is built on PyTorch 2.0.1 with CUDA 12.1 support for GPU-accelerated computation. The torchvision library provides the pretrained

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

ResNet-50 model weights and standard image transformation utilities. The EfficientNet-B4 implementation and pretrained weights are loaded from the timm library. The pytorch-grad-cam library implements the Grad-CAM computation. The web application backend is implemented using the Flask framework, which exposes REST API endpoints for image classification, result retrieval, and expert feedback submission. The frontend user interface is built with React 18 and communicates with the backend via asynchronous HTTP requests.

The training dataset comprises 43,000 images split across five product categories with an equal balance of authentic and counterfeit examples. The dataset is divided into training, validation, and test subsets in an eighty-ten-ten proportion using stratified sampling. Both backbone models are fine-tuned using the AdamW optimizer. The ResNet-50 backbone is trained with a learning rate of one times ten to the negative three for the classification head and one times ten to the negative four for the convolutional backbone over thirty epochs with cosine annealing scheduling. The EfficientNet-B4 backbone is trained with a learning rate of five times ten to the negative four and a linear warmup of five epochs followed by cosine decay over twenty-five epochs. Training is conducted on a single NVIDIA A100 GPU with a batch size of sixty-four. Augmentation strategies including random horizontal flipping, rotation up to thirty degrees, colour jitter with brightness and contrast variation of twenty percent, and MixUp with alpha of 0.2 are applied during training to improve generalization.

The deployment configuration uses a Kubernetes cluster with a single GPU node for the inference service and a CPU node for the Flask web application and PostgreSQL database. The Flask API enforces JWT-based authentication for all endpoints and rate limits of sixty requests per minute per authenticated client. The feedback database stores corrective labels along with references to the original image objects stored in an encrypted S3-compatible object store. All inter-service communication uses HTTPS with TLS 1.3. The Docker image is approximately 6 gigabytes including CUDA runtime libraries and PyTorch dependencies.

## EXPERIMENTAL EVALUATION

The Counterfeit Product Detector was evaluated on a 4,300-image held-out test set constructed from the ten percent test split of the full 43,000-image dataset, maintaining equal representation of authentic and counterfeit examples across all five product categories. The primary evaluation metric is classification accuracy, defined as the proportion of test images correctly classified as authentic or counterfeit. Secondary metrics include precision, recall, F1-score for the counterfeit class, area under the receiver operating characteristic curve, and Expected Calibration Error for confidence quality assessment.

The ResNet-50 backbone alone achieved a test accuracy of 93.1 percent on the full test set. The EfficientNet-B4 backbone alone achieved 93.8 percent accuracy. The ensemble of the two models achieved 95.7 percent overall accuracy, demonstrating that the two backbone architectures capture complementary visual features that are not fully correlated. The improvement from 93.8 to 95.7 percent represents a meaningful reduction in error rate that translates to a significant number of additional correct classifications at operational scale. Per-category accuracy ranged from 93.4 percent for packaged food products, the most visually heterogeneous category, to 97.8 percent for pharmaceutical tablets, where counterfeit examples consistently exhibit measurable deviations in colour uniformity and surface texture relative to the authentic standard.

Confidence calibration was evaluated by measuring the Expected Calibration Error before and after temperature scaling. The pre-calibration ECE for the ResNet-50 backbone was 0.087, indicating significant overconfidence in its probability estimates. After temperature scaling with an optimized temperature parameter of 1.47, the post-calibration ECE was reduced to 0.021. The EfficientNet-B4 backbone showed a corresponding improvement from 0.074 to 0.018. The ensemble achieved a final ECE of 0.016, confirming that the reported confidence scores accurately reflect empirical accuracy across the full confidence range. This calibration quality is essential for the two-tier decision protocol to correctly identify the borderline predictions that require expert review.

The two-tier decision pipeline routed 312 of 4,300 test images, representing 7.3 percent of all predictions, to the expert review queue due to confidence scores in the uncertainty band between 0.40 and 0.60. Expert reviewers correctly classified 306 of these 312 borderline cases, achieving 98.1 percent accuracy on the uncertain subset. The combined system accuracy including expert review was 96.8 percent, compared to 95.7 percent for the automated system alone, confirming that the hybrid approach adds meaningful value for the most challenging authentication cases.

Robustness testing evaluated the system under degraded input conditions including JPEG compression at quality factors between 90 and 30, uniform illumination reduction to 70, 50, and 30 percent of standard levels, and rectangular occlusions covering 10, 20, and 30 percent of image area. Classification accuracy remained above 93 percent for JPEG quality factors as low as 50 and illumination levels at 50 percent of standard. Accuracy degraded

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

to 87.3 percent at 30 percent illumination and 84.1 percent at 30 percent occlusion, identifying low-light photography and heavily obscured product labels as the primary failure conditions for future model improvement.
Figure 4: Evaluation Results — Ensemble model achieves 95.7% accuracy across five categories; calibrated ECE of 0.016; expert review raises combined accuracy to 96.8%

## CONCLUSION

The Counterfeit Product Detector using Deep Learning presented in this paper demonstrates how modern computer vision and deep learning techniques can be applied to address the significant and growing problem of counterfeit product proliferation in global markets. The system integrates transfer learning from pretrained ResNet-50 and EfficientNet-B4 convolutional neural network backbones, ensemble classification, temperature-scaled confidence calibration, Grad-CAM visual explanation, a two-tier automated and expert review decision pipeline, and a continuous learning mechanism driven by expert feedback. These components work together to deliver a product authentication system that is accurate, interpretable, scalable, and continuously improvable.

The experimental evaluation demonstrates that the ensemble model achieves 95.7 percent classification accuracy across five diverse product categories on a 4,300-image test set, with the two-tier decision protocol raising the combined accuracy to 96.8 percent when human expert review is applied to borderline cases. Confidence calibration ensures that the reported probability scores accurately reflect empirical accuracy, enabling reliable uncertainty quantification. Robustness testing confirms that the system maintains strong performance under realistic imaging degradation conditions including JPEG compression and reduced illumination.

The modular pipeline architecture ensures that individual components can be upgraded as deep learning research advances, and the Kubernetes-based deployment infrastructure enables horizontal scaling to accommodate growing operational demand. The continuous learning pipeline ensures that the system improves over time as expert feedback accumulates, maintaining its effectiveness against evolving counterfeit methods. Future work will explore vision transformer architectures for improved fine-grained feature discrimination, multi-modal authentication combining visual analysis with NFC chip verification, lightweight on-device inference for consumer mobile applications, and zero-shot authentication for product categories not present in the training dataset using large-scale vision-language model embeddings.

## REFERENCES

1) Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” in Proc. NeurIPS, Lake Tahoe, NV, USA, 2012, pp. 1097–1105.
2) K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” in Proc. ICLR, San Diego, CA, USA, 2015.
3) K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” in Proc. IEEE/CVF CVPR, Las Vegas, NV, USA, 2016, pp. 770–778.
4) M. Tan and Q. V. Le, “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks,” in Proc. ICML, Long Beach, CA, USA, 2019, pp. 6105–6114.
5) R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, “Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization,” in Proc. IEEE/CVF ICCV, Venice, Italy, 2017, pp. 618–626.
6) Guo, G. Pleiss, Y. Sun, and K. Q. Weinberger, “On Calibration of Modern Neural Networks,” in Proc. ICML, Sydney, Australia, 2017, pp. 1321–1330.
7) Gao, Y. Lin, and X. Ma, “Siamese Networks for Product Authentication in E-Commerce,” in Proc. IEEE ICME, Shanghai, China, 2019, pp. 1–6.
8) H. Zhang, M. Cisse, Y. N. Dauphin, and D. Lopez-Paz, “MixUp: Beyond Empirical Risk Minimization,” in Proc. ICLR, Vancouver, Canada, 2018.
9) S. Jain and A. Garg, “Pharmaceutical Tablet Authentication Using Deep Convolutional Neural Networks,” J. Pharm. Biomed. Anal., vol. 189, p. 113452, Sep. 2020.
10) Dosovitskiy et al., “An Image is Worth 16×16 Words: Transformers for Image Recognition at Scale,” in Proc. ICLR, Virtual, 2021.