Volume 11, Issue 5, May – 2026 International Journal of Innovative Science and Research Technology
ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26may998

AuthentiCheck: A Multimodal AI Framework for Counterfeit Detection

M. Shaik Suzain Saba[1]; Dr. Girish Kumar D.[2]; Sreelakshmi J.[3]

[1]PG Student, Department of MCA, Ballari Institute of Technology & Management, Ballari.
[2]Professor and HOD, Department of MCA, Ballari Institute of Technology & Management, Ballari.
[3]Assistant Professor, Department of MCA, Ballari Institute of Technology & Management, Ballari.

Publication Date: 2026/06/06

Abstract: Accelerated growth in online retail has sharply elevated the volume of fake merchandise circulating across digital markets, creating wide-ranging economic, legal, and safety consequences for both consumers and brand owners. Counterfeit goods are typically engineered to closely mimic genuine products in appearance, making human-led verification unreliable and difficult to scale. Conventional rule-based detection tools have consistently struggled to keep pace with sophisticated and continuously evolving forgery methods. AuthentiCheck is introduced as an AI-powered detection framework that unifies computer vision, NLP, OCR, and metadata validation within a single pipeline. The system jointly analyzes product imagery, listing text, seller profiles, and pricing signals. A cross-modal fusion engine synthesizes evidence from all channels to generate an interpretable authenticity rating and risk classification. Empirical testing confirms this multi-source strategy significantly outperforms single-channel alternatives. The framework prioritizes scalability, interpretability, and ethical deployment.

Keywords: *Counterfeit Detection, Computer Vision, Natural Language Processing, Multi-Modal AI, E-Commerce Security.*

How to Cite: M. Shaik Suzain Saba; Dr. Girish Kumar D.; Sreelakshmi J. (2026) AuthentiCheck: A Multimodal AI Framework for Counterfeit Detection. *International Journal of Innovative Science and Research Technology*, 11(5), 3451-3457. https://doi.org/10.38124/ijisrt/26may998

I. INTRODUCTION

The proliferation of online retail platforms has fundamentally reshaped how consumers discover and purchase goods, granting access to an extensive range of products across geographic boundaries. While this transition has generated substantial economic value, it has equally enabled the widespread circulation of counterfeit merchandise. Fake electronics, luxury goods, cosmetics, pharmaceuticals, and apparel are routinely listed on major platforms, frequently indistinguishable from genuine items through visual inspection alone.

The harms caused by counterfeit goods extend well beyond financial losses for manufacturers — they degrade brand trust and, in sensitive product categories such as medications, pose direct health risks to consumers. Manual review processes cannot operate at the speed or scale required by modern e-commerce, while fixed-rule automated systems lose effectiveness as fraudulent tactics evolve. This creates a pressing demand for adaptive, intelligent solutions capable of processing heterogeneous data and flagging authenticity anomalies with high precision.

Recent advances in artificial intelligence provide promising tools for addressing this challenge. Computer vision enables fine-grained visual comparison of packaging and logos, while NLP models can surface characteristic language signatures embedded in deceptive product listings. Single-modality systems, however, capture only a fraction of the available authenticity signals.

AuthentiCheck closes this gap through a unified architecture that concurrently leverages visual, textual, and metadata evidence to deliver more dependable authenticity assessments.

➢ *The Contributions of this Work are as Follows:*

- Design of a scalable, multi-modal counterfeit detection architecture.
- Integration of vision, language, OCR, and metadata analysis into a unified pipeline.
- Development of a fusion-based authenticity scoring mechanism.
- Emphasis on interpretability and actionable reporting for end users.

II. BACKGROUND & MOTIVATION

Counterfeit production methods have grown substantially more sophisticated in tandem with advances in manufacturing, precision printing, and global supply chains. Modern fake goods are no longer crude imitations; producers now invest in high-fidelity packaging reproduction, accurate

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26may998

logo duplication, and branding materials nearly identical to those of genuine manufacturers. As a result, even well-informed buyers and trained platform reviewers struggle to distinguish fake products from authentic ones through visual examination alone.

Digital commerce environments intensify this difficulty because buyers have no opportunity to physically examine merchandise before purchasing. All judgment must rely on digital artifacts product images, listing text, pricing details, and seller account data. Bad actors systematically exploit this reliance by staging misleading photographs, overstating claimed specifications, and using persuasive language crafted to manufacture false credibility. These tactics routinely bypass conventional fraud filters built on static keyword lists or predefined rule sets.

From an information analytics perspective, authenticity signals are distributed across multiple independent data channels. Visual content encodes packaging fidelity and brand accuracy; text carries semantic and stylistic cues; metadata — covering seller history, account age, price benchmarks, and shipping origin — supplies contextual evidence unavailable from images or text alone. Systems restricted to a single channel will necessarily yield incomplete assessments.

AuthentiCheck is motivated by the opportunity to close these gaps through concurrent multi-source analysis. By processing all available data channels in parallel, the framework can identify cross-modal correlations that remain invisible when channels are evaluated independently. This comprehensive strategy improves detection reliability while aligning with the operational demands of production ecommerce environments, where decisions must be fast, auditable, and adaptable to shifting fraud patterns.

III. RELATED WORK

Early counterfeit detection work depended on deterministic rule engines and domain expert heuristics. These approaches typically flagged listings matching banned keyword lists, known fraudulent seller identifiers, or manually compiled image templates. While serviceable in narrow, controlled settings, they lacked the adaptability needed to remain effective as deception tactics diversified.

Later research introduced data-driven methods, particularly image classifiers, for detecting packaging defects, logo irregularities, and visual markers associated with fake merchandise. Deep convolutional networks achieved strong classification accuracy and enabled automated identification of visually atypical product listings.

Image-only evidence is inherently limited, however, since premium-quality fakes are deliberately engineered to be visually indistinguishable from genuine goods.

A concurrent research stream applied NLP to identify deceptive listing text and fabricated customer reviews.

Sentiment classifiers, syntactic detectors, and pre-trained language models each demonstrated capacity to surface writing patterns correlated with fraudulent intent. Text-only systems nevertheless remain susceptible to intentional stylistic manipulation, constraining their practical robustness.

Despite these capabilities, text-only systems remain susceptible to intentional stylistic manipulation and adversarial prompt engineering, which constrains their practical robustness.

Growing research interest in multi-source learning has yielded promising results across related domains including synthetic media detection, biometric verification, and financial fraud identification. These studies consistently demonstrate that classifiers fusing complementary input channels outperform single-modality alternatives.

Despite this progress, most multi-modal counterfeit detection systems remain at the prototype stage, with limited treatment of integration challenges, interpretability requirements, and scalability constraints. The present work contributes by delivering a fully integrated, end-to end system built with practical deployment as an explicit objective.

➢ *System Architecture Accuracy Metrics*

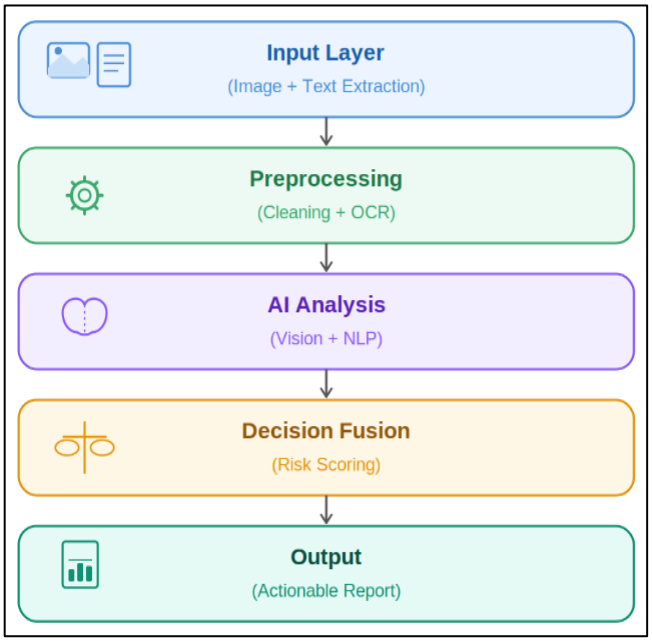

Fig 1 Overall System Architecture of the AuthentiCheck Framework

AuthentiCheck is organized as a stratified pipeline. The intake layer ingests product images, listing text, pricing data, and seller profile attributes. A preprocessing stage normalizes inputs via image standardization, OCR-based text extraction, and NLP cleaning. Three parallel analytical modules process visual, linguistic, and metadata signals respectively. Their outputs converge in a central scoring engine producing a unified authenticity rating. Results are delivered through an interactive reporting interface.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26may998

IV. MULTI-MODAL DATA PROCESSING

➢ *Visual Analysis*

Each product image is inspected for logo placement accuracy, chromatic fidelity, typographic consistency, structural packaging symmetry, and overall surface finish quality. Deep convolutional networks generate high-dimensional visual embeddings that are matched against curated brand reference libraries maintained for each monitored product category.

➢ *Textual Analysis*

NLP pipelines process listing descriptions and associated customer reviews to identify deceptive content. Key risk signals include atypical grammar, exaggerated promotional language, semantic contradictions within the same listing, and keyword patterns empirically associated with counterfeit product postings associated with counterfeit listings.

➢ *OCR and Label Verification*

An OCR engine recovers textual content embedded in product label imagery — including batch identifiers, provenance details, and barcode data. Recovered strings are cross-referenced against established formatting standards and verified brand records to flag anomalies indicating non-genuine origin.

➢ *Metadata Validation*

Seller-level signals — including account standing, historical ratings, listing age, price deviation relative to brand benchmarks, and shipping geography — are systematically assessed. Listings showing anomalous pricing or originating from atypical seller profiles receive elevated risk scores.

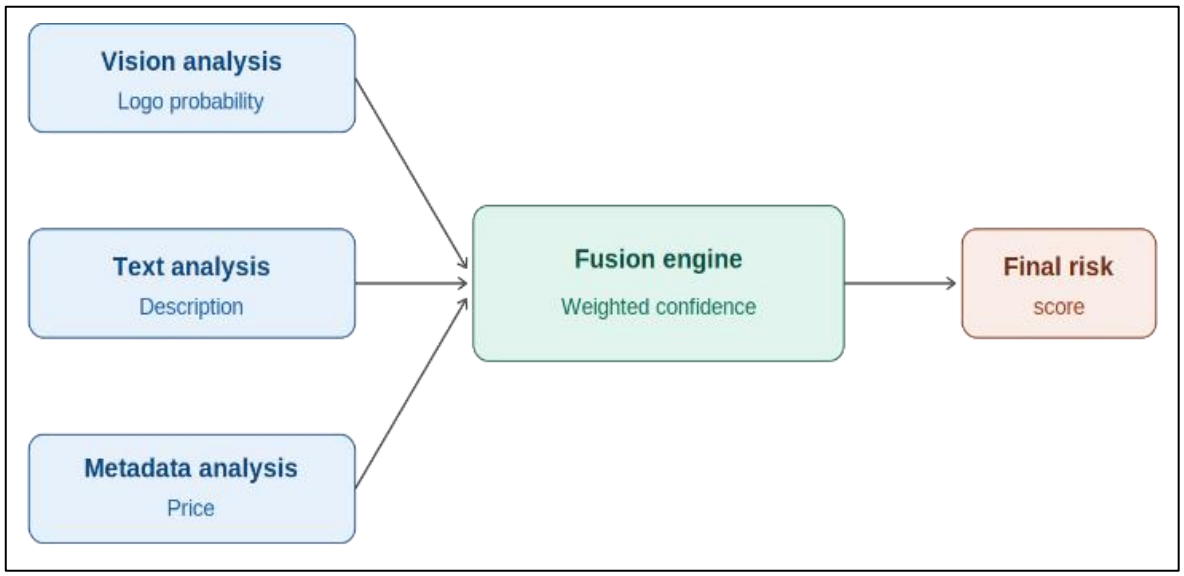

Fig 2 Fusion Based Decision Module for Authenticity Scoring.

V. DECISION FUSION & SCORING

The decision fusion engine forms the central reasoning component of AuthentiCheck. Because each input modality contributes fundamentally different signal types, naive threshold-based combination fails to capture the nuanced interdependencies among channels. Instead, a weighted aggregation approach is applied, combining calibrated probability scores from each analytical module according to empirically determined weights.

Every analytical module produces a scalar probability representing the estimated likelihood of counterfeit origin for a given listing. These scores are normalized to a common scale prior to aggregation, ensuring no single channel disproportionately skews the combined result. Module weights are calibrated through validation experiments reflecting each channel's measured contribution to overall detection accuracy.

The weighted aggregate is then discretized into interpretable risk tiers that support operational decisionmaking. Mapping continuous scores to categorical labels reduces cognitive burden for non-specialist reviewers while preserving the probabilistic foundation of the assessment. Crucially, the fusion layer maintains a complete audit trail linking every verdict to its contributing analytical signals. This design embeds accountability throughout the pipeline and supports regulatory compliance requirements in commercial deployment environments.

Each analytical component contributes a weighted probability score to the fusion layer. The aggregation logic ensures balanced contribution from all input channels, preventing any single modality from dominating the outcome. The composite score is mapped to one of three risk categories: Authentic, Suspicious, or High-Risk Counterfeit categories such as *Authentic*, *Suspicious*, or *High-Risk Counterfeit*.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26may998

➢ *Dashboard and Reporting*

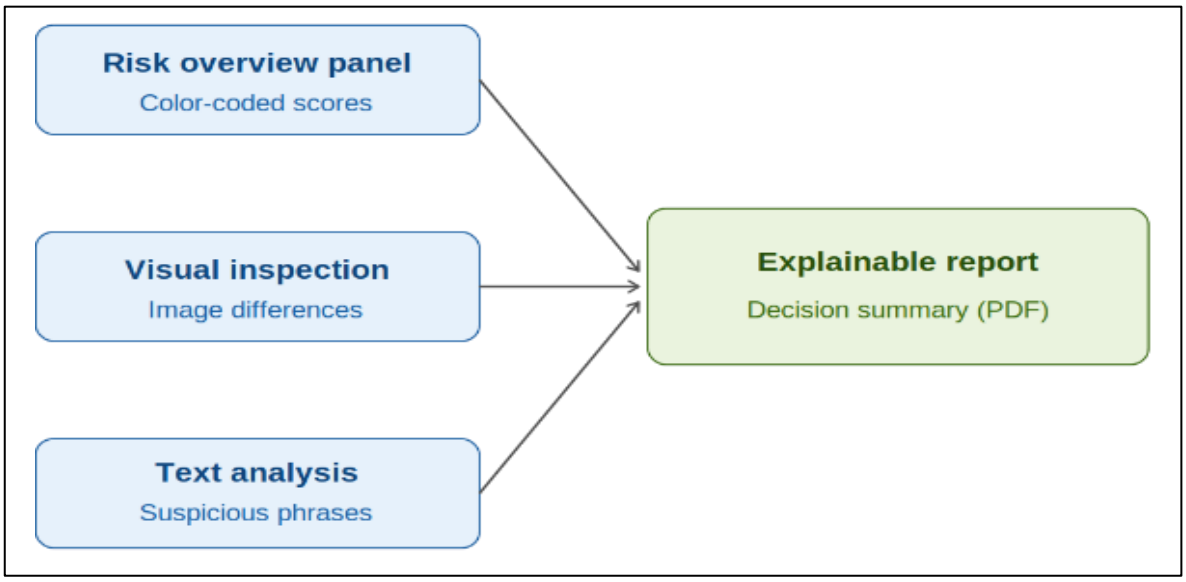

Fig 3 User-Facing Analytics Dashboard for Counterfeit Detection Results.

The reporting interface consolidates authenticity ratings, image-level anomaly annotations, linguistic risk indicators, and seller credibility summaries within a unified view. Users can explore historical detection trends, review evidence supporting individual verdicts, and export structured verification reports.

VI. EXPERIMENTAL EVALUATION

System performance was evaluated against a heterogeneous corpus comprising both structured benchmark data and organic product listings drawn from live e-commerce platforms. The evaluation set spanned multiple merchandise categories — electronics, apparel, cosmetics, and consumer goods — to ensure broad representativeness across product types.

AuthentiCheck was compared against three singlechannel reference systems, each restricted to one input modality. Classification performance was quantified using accuracy, precision, recall, and adversarial robustness measures. Evaluation scenarios specifically targeted difficult cases where counterfeit listings were engineered to closely replicate authentic products.

AuthentiCheck delivered superior outcomes across all test conditions relative to every single-channel baseline. The multi-source fusion approach also showed markedly stronger stability when individual data channels were deliberately manipulated, validating the hypothesis that concurrent multi-channel analysis yields more dependable counterfeit detection.

➢ *Dataset Description*

The experimental dataset comprised 24,590 product assembled from publicly accessible e-commerce data sources and curated academic repositories. Five merchandise categories were covered: electronics, apparel, cosmetics, consumer goods, and pharmaceuticals. All listings were independently labeled by domain specialists and crossvalidated against verified brand reference data to ensure annotation integrity. Table 1 presents the category-level distribution.

Table 1 Dataset Distribution by Product Category

Product Category	Authentic Samples	Counterfeit Samples	Total Listings
Electronics	3,200	3,150	6,350
Apparel	2,800	2,760	5,560
Cosmetics	2,500	2,480	4,980
Consumer Goods	2,100	2,050	4,150
Pharmaceuticals	1,800	1,750	3,550
Total	12,400	12,190	24,590

- *Evaluation Metrics and Baseline Comparison*

AuthentiCheck was benchmarked against three single-modality baselines: Vision-Only (convolutional feature extraction), Text-Only (transformer-based NLP), and Metadata-Only (gradient-boosted classifier). All models were trained and evaluated on identical dataset splits using 80/10/10 train/validation/test partitions with stratified sampling. Table 2 presents the aggregated performance across all categories.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26may998

Table 2 Overall Performance Comparison Across Methods

Method	Accuracy (%)	Precision (%)	Recall (%)	F1- Score (%)
Vision-Only	82.3	80.1	79.4	79.7
Text-Only	78.6	76.8	77.2	77
MetadataOnly	74.2	72.5	73.1	72.8
AuthentiCh eck (MultiModal)	94.7	93.8	94.1	93.9

AuthentiCheck surpassed all three reference systems, recording 94.7% accuracy and a 93.9% F1-score. The 12.4 percentage-point accuracy improvement over Vision-Only (82.3%) quantifies the tangible benefit of integrating multiple data channels. The Metadata-Only classifier posted the weakest results overall, confirming that pricing and seller attributes alone are insufficient as standalone discriminative features.

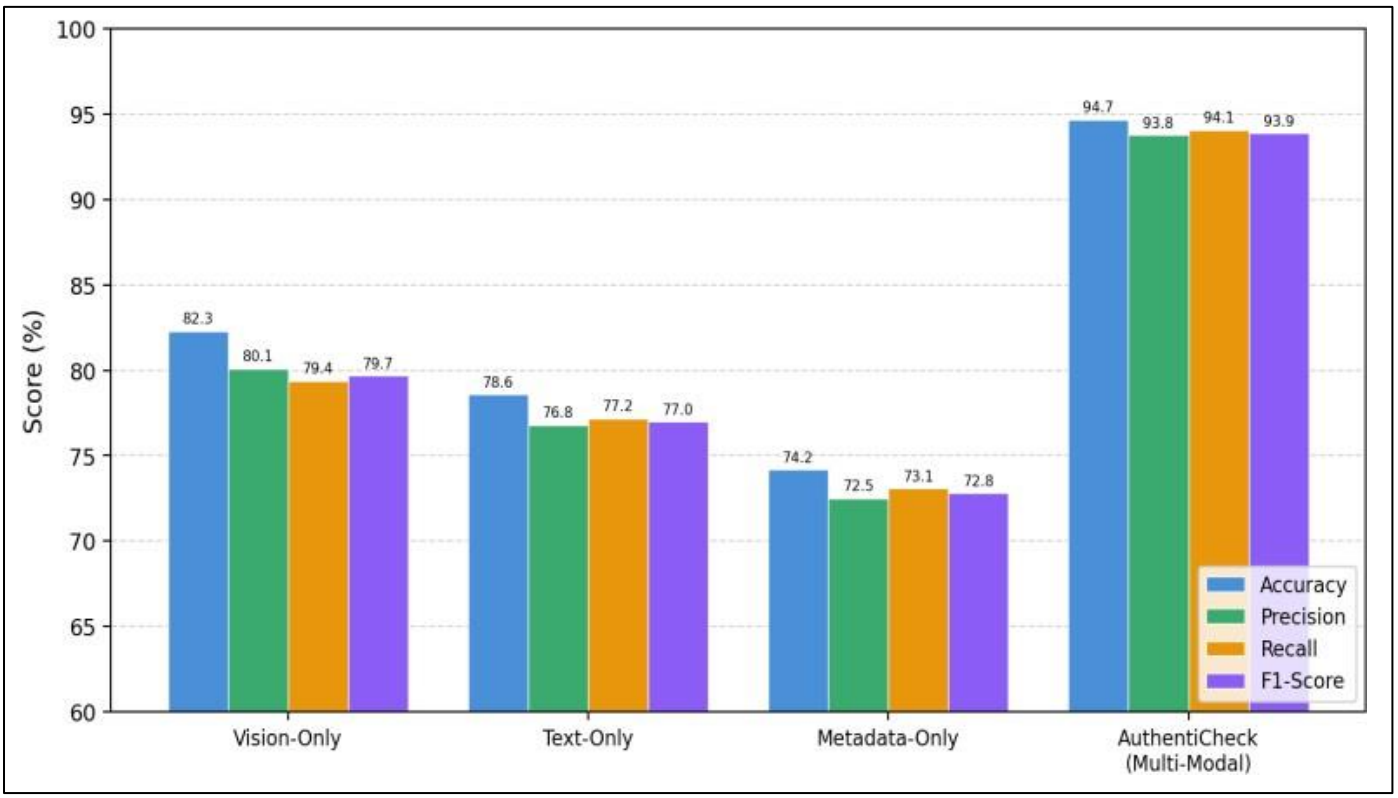

Fig 4 Performance Comparison Across Methods: Accuracy, Precision, Recall, and F1-Score (%).

- *Per-Category Detection Performance*

Disaggregated detection figures are shown in Table 3. Electronics and consumer goods achieved the highest accuracy at 95.2% and 95.1% respectively, attributable to rich brand imagery availability. Pharmaceutical listings were most challenging due to strong visual and textual overlap between genuine and counterfeit samples; AuthentiCheck nonetheless sustained 92.4% accuracy.

Table 3 Per-Category Detection Performance of AuthentiCheck

Category	Accura cy (%)	Precision (%)	Recall (%)	F1-Score (%)
Electronics	95.2	94.6	95.0	94.8
Apparel	93.8	92.9	93.5	93.2
Cosmetics	94.6	93.7	94.3	94.0
Consumer Goods	95.1	94.2	94.8	94.5
Pharmaceuticals	92.4	91.8	92.1	91.9

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26may998

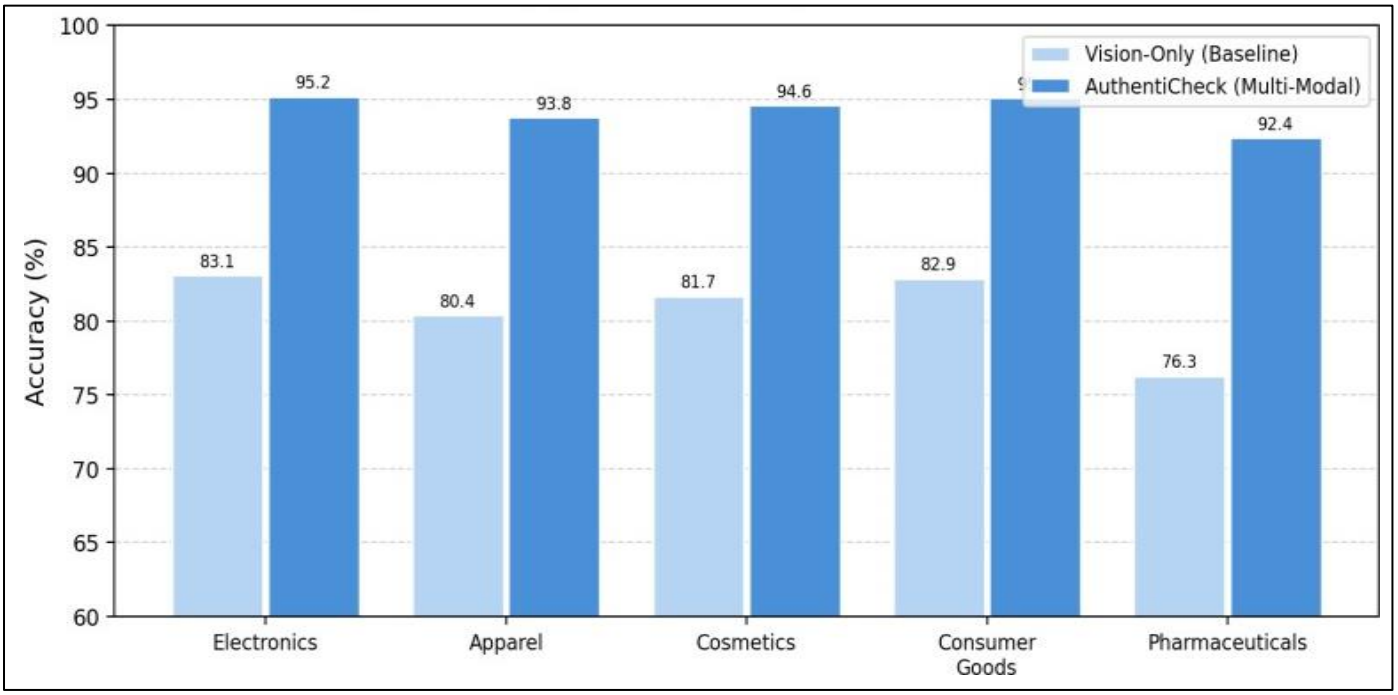

Fig 5 Per-Category Detection Accuracy: Vision-Only Baseline vs AuthentiCheck Multi-Modal Framework.

➢ *ROC Analysis and AUC Scores*

Threshold-level discriminative performance was characterized through ROC analysis. AuthentiCheck attained an AUC of 0.978 — a 7.7-point gain over Vision-Only (AUC 0.901) and a 10.2-point gain over Text-Only (AUC 0.876). The ROC profile confirms that the fusion model sustains high true positive rates even at low false positive thresholds, an important property in operational contexts where spurious alerts impose meaningful costs.

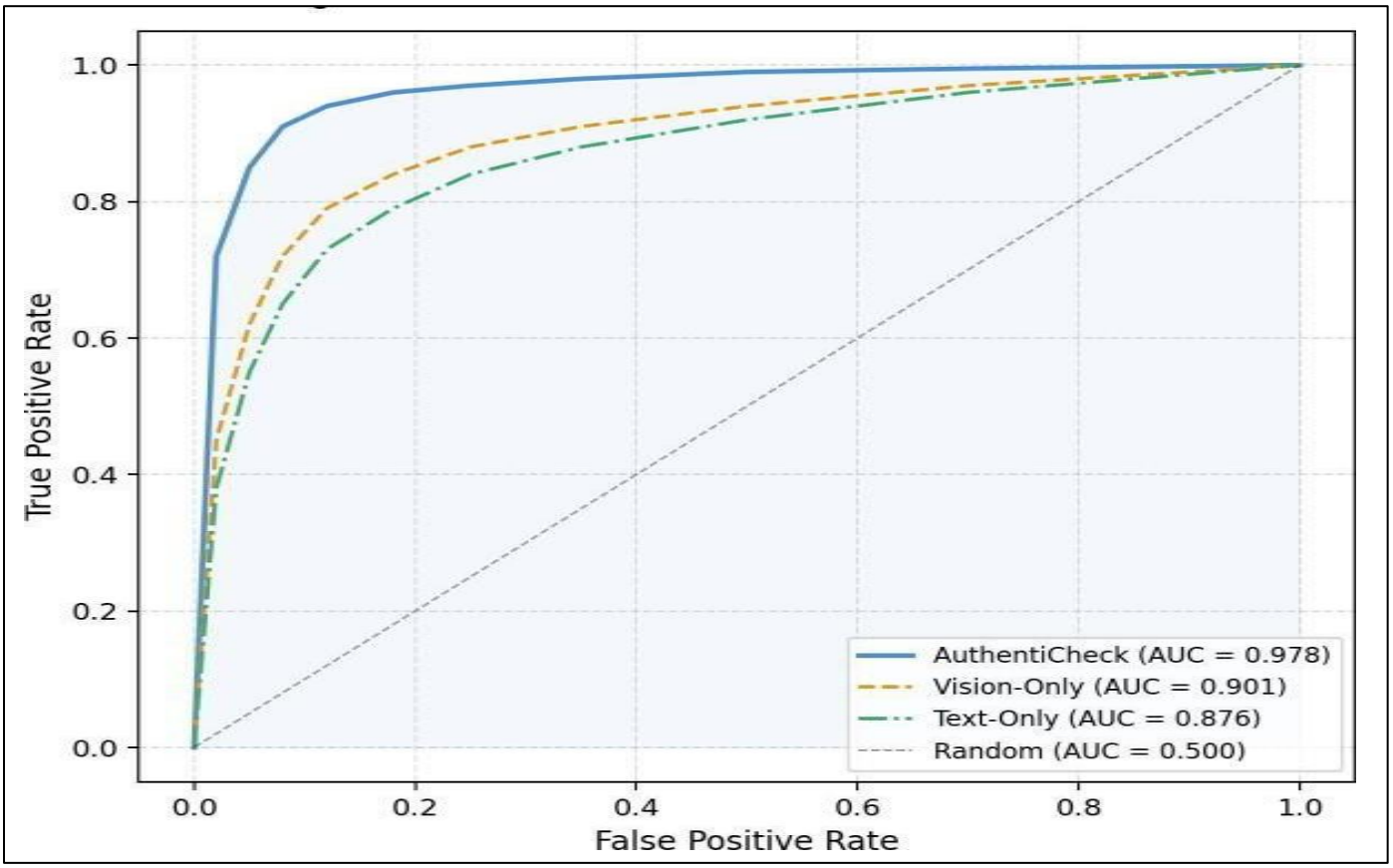

Fig 6 ROC Curves Comparing AuthentiCheck against Singlemodality Baseline Methods.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26may998

➢ *Adversarial Robustness*

Adversarial robustness was assessed on 1,200 synthetic listings designed to exploit individual data channels. When authentic-looking images were combined with fabricated descriptions, Vision-Only accuracy fell to 61.3% while AuthentiCheck retained 89.7% by drawing on textual and metadata signals. In complementary tests with manipulated metadata alongside genuine imagery and text, Metadata-Only accuracy collapsed to 48.2% while AuthentiCheck held at 88.4%. Cross-channel fusion demonstrably hardens the system against targeted singlemodality attacks.

VII. ETHICAL AND GOVERNANCES

➢ *Considerations*

Introducing automated detection tools into commercial environments raises substantive ethical and governance responsibilities. Misclassification can unjustly penalize legitimate sellers or restrict consumer access to genuine goods. To address these risks, AuthentiCheck is architected as an advisory decision-support system that augments human judgment rather than replacing it with autonomous enforcement actions.

Accountability is maintained through explanatory outputs that attribute each risk verdict to specific contributing signals across modalities. This enables platform administrators and affected sellers to review, understand, and contest flagged decisions when warranted. The system additionally avoids ingesting protected demographic attributes and applies data minimization principles throughout.

Regular fairness audits and performance monitoring are integral to system governance. These measures ensure consistent behavior across demographic and geographic segments while maintaining compliance with data protection regulations. Ethical deployment considerations are embedded throughout the system lifecycle, from data.

VIII. LIMITATION AND FUTURE WORK

Current constraints include sensitivity to the completeness of reference brand datasets and significant compute overhead associated with processing high-resolution imagery at scale. Planned future directions include blockchain-anchored product certification for tamper-evident provenance tracking, expanded multilingual NLP support to broaden geographic coverage, and model compression techniques targeting edge-device deployment.

IX. CONCLUSION

AuthentiCheck has been presented as a holistic, multi-modal AI framework designed to identify counterfeit product listings in ecommerce environments. By unifying image analysis, text understanding, OCR-based label verification, and metadata assessment, the system delivers robust and interpretable authenticity scoring. The fusion-based architecture directly addresses the documented limitations of channel-specific detection methods, with empirical results demonstrating substantially improved reliability across diverse product categories and adversarial scenarios. Interpretability, fairness, and user agency have been embedded as first-class design requirements throughout.

AuthentiCheck carts a viable path toward scalable, trustworthy counterfeit detection and holds meaningful potential for adoption across modern digital commerce ecosystems.

REFERENCES

[1]. R. Baeza-Yates, “Bias on the web and algorithmic systems,” Communi- cations of the ACM, vol. 61, no. 6, pp. 54–61, 2018.
[2]. D. Marler and J. W. Boudreau, “An evidence-based review of human resource analytics,” International Journal of Human Resource Manage- ment, vol. 28, no. 1, pp. 3–26, 2017.
[3]. S. Lundberg and S.-I. Lee, “A unified framework for interpreting predictions of complex machine learning models,” in Advances in Neural Information Processing Systems (NeurIPS), 2017, pp. 4765–4774.
[4]. J. Angrave, A. Charlwood, I. Kirkpatrick, M. Lawrence, and M. Stuart, “HR analytics and big data: Transforming decision- making in human resource management,” Human Resource Management Journal, vol. 26, no. 1, pp. 1–13, 2016.
[5]. S. Barocas and A. D. Selbst, “Big data’s disparate impact on societal decision systems,” California Law Review, vol. 104, no. 3, pp. 671–732, 2016.
[6]. I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. MIT Press, 2016.
[7]. A. Vaswani et al., “Attention is all you need,” in Advances in Neural Information Processing Systems (NeurIPS), 2017, pp. 5998–6008.
[8]. Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, pp. 436–444, 2015.
[9]. G. James, D. Witten, T. Hastie, and R. Tibshirani, An Introduction to Statistical Learning with Applications in R. Springer, 2013.
[10]. A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, Bayesian Data Analysis, 3rd ed. CRC Press, 2013.
[11]. J. Han, M. Kamber, and J. Pei, Data Mining: Concepts and Techniques, 3rd ed. Elsevier, 2012.
[12]. J. Pearl, Causality: Models, Reasoning, and Inference, 2nd ed. Cam- bridge University Press, 2009.
[13]. T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning, 2nd ed. Springer, 2009.
[14]. L. Breiman, “Random forests,” Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.
[15]. M. Stone, “Cross-validatory choice and assessment of statistical predic- tions,” Journal of the Royal Statistical Society, Series B, vol. 36, no. 2, pp. 111–147, 1974.