

Multimodal Counterfeit Intelligence Framework for Authenticity Verification in Online Market Places

Syed Jasmin¹, C Rajesh², K Vishnupriya³, K M Doraswamy⁴, N Rohith Kumar⁵

^{1,2,3,4,5} *Computer Science and Information Technology, Siddharth Institute of Engineering & Technology*

ABSTRACT - This project addresses the erosion of consumer trust caused by counterfeit e-commerce listings by developing a multi-modal detection system that identifies inconsistencies across textual descriptions, product images, and seller metadata. The technical framework utilizes transformer-based embeddings to flag linguistic anomalies and CNN/ViT encoders to compare listing images against authentic brand references. By integrating these features through both early and late fusion strategies, the system generates a risk-prioritized likelihood score accompanied by explanatory highlights for human reviewers. Validated against adversarial examples and evaluated using metrics such as F1-score and false-positive rates, the solution supports scalable deployment through batch and streaming APIs. The resulting interactive dashboard provides a robust defense mechanism for online marketplaces.

Keywords: Multi-Modal Deep Learning, Fake Product Detection, Transformer Models, CNN (convolutional Neural Networks), Vision Transformers (ViT), Feature Fusion, Seller Metadata Analysis.

1. INTRODUCTION

While online shopping offers unmatched convenience, the rise of sophisticated fake listings has outpaced traditional detection methods, causing financial loss and eroding platform trust. Modern fraud often bypasses simple text-based filters by manipulating images or pricing, highlighting the need for an intelligent system that can analyze product data with the nuance of a human reviewer. This project introduces a Multi-Modal Fake Product Detection System that unifies text, images, and seller metadata into a single decision-making process. By leveraging Transformer-based models and Vision Transformers through early and late fusion, the system identifies subtle inconsistencies such as visual-textual mismatches or abnormal pricing to provide real-time risk scores and actionable insights for e-commerce security

2. SYSTEM ANALYSIS

Existing System:

Current product detection techniques include:

- **Only text-based analysis:** Ignores visual and behavioral data, missing sophisticated cross-modal fraud.
- **Manual human review:** Too slow and costly to handle the massive volume of daily listings.
- **Image verification is limited:** Fails to detect visual manipulations or mismatches with product claims.
- **False positives and negatives:** Causes loss of revenue from genuine sellers or fails to block scammers.
- **Not deeply analyzed:** Lacks the depth needed to uncover complex, hidden relationships in the data.

Limitations of Existing Systems:

- Relies heavily on single-modal analysis, mainly text.
- Fails to detect mismatches between product images and descriptions.
- High dependency on manual review increases time and cost.
- Limited analysis of seller behaviour and pricing anomalies.
- High chances of false positives and false negatives.
- Not scalable for large-scale, real-time e-commerce environments.

3. PROPOSED SYSTEM

Support Vector Machine (SVM) is a supervised learning algorithm used for classification and regression tasks. It works by finding an optimal hyperplane that separates two classes with maximum margin.

- **Multi-modal approach:** Combines text, images, and metadata to ensure a comprehensive evaluation of every product listing.
- **Transformer model:** Extracts deep semantic meaning from product descriptions to detect deceptive language and exaggerated claims.
- **CNN / Vision Transformer (ViT):** Analyzes visual features to identify discrepancies between product photos and their stated categories.

- **Analyze seller metadata:** Evaluates account history and pricing patterns to distinguish legitimate businesses from high-risk accounts.
- **Perform early fusion and late fusion:** Merges data at different stages to capture complex relationships between various information sources.
- **Generate risk score:** Provides a numerical probability of fraud to help platforms prioritize high-risk listings for immediate action.

Build a dashboard: Offers an intuitive interface for reviewers to inspect flagged products and understand the model's reasoning.

Advantages of Proposed System:

- Uses combined text, image, and metadata analysis for higher accuracy.
- Detects image–text mismatches effectively
- Identifies risky sellers using behavioural patterns.
- Provides explainable outputs for better trust and review.
- Supports real-time detection through API integration.
- Scalable for large e-commerce platforms.

EXPERIMENTAL RESULTS:

The proposed Multimodal system achieves:

- High detection accuracy
- Bait-and-Switch Detection
- Reduced false positive rate
- Faster detection time compared to content-based methods
- Interactive Dashboard

Diagram 1: System Architecture

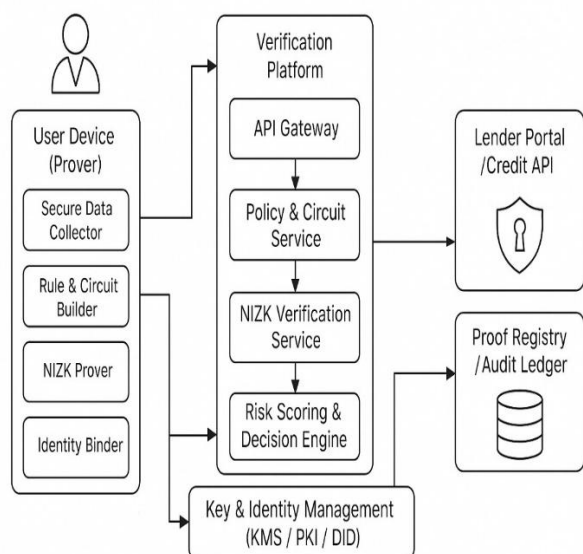


Fig -1: Figure

4. CONCLUSION

This project presents an effective and scalable Fake Product Detection System that addresses the growing challenge of counterfeit and misleading product listings in online marketplaces. By combining text analysis, image analysis, and seller metadata evaluation, the system provides a holistic approach to identifying suspicious products that traditional single-modal methods often fail to detect.

The proposed system improves detection accuracy through multi-modal fusion and enhances transparency by generating clear risk scores and explanations. Its modular, API-driven design makes it suitable for real-world deployment on large e-commerce platforms. Overall, the system contributes to building safer online shopping environments by protecting consumers, supporting genuine sellers, and strengthening platform trust.

FUTURE SCOPE

In the future, the system will be enhanced to handle large-scale, real-time product streams by optimizing model inference and incorporating graph-based analysis to uncover hidden networks of fraudulent sellers. By integrating **continual and** semi-supervised learning, the model will automatically adapt to emerging fraud patterns based on reviewer feedback, while advanced multimodal transformers will improve the detection of sophisticated image manipulations. Additionally, the scope includes cross-platform data sharing and the adaptation of the system for mobile environments and evolving global regulations, ensuring a robust, future-ready defense against e-commerce fraud.

REFERENCES

1. Liang, Y., et al. (2022). "Multimodal false information detection method based on Text-CNN and SE module." PLoS ONE.
2. Gowda, M. R., & Pradeep, K. (2025). "Deepfake Detection Using CNN and CviT Transformers." IJCRT.
3. Dosovitskiy, A., et al. (2021). "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale." ICLR. (The original paper for Vision Transformers - ViT).
4. Devlin, J., et al. (2019). "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." NAACL.

BIOGRAPHY



I, **SYED JASMIN**, currently working as an Assistant Professor in the Department of Computer Science and Information Technology at Siddharth Institute of Engineering & Technology, Puttur, Andhra Pradesh, India. I am having 3 years of teaching experience in engineering education. I am pursuing my Ph.D. in Computer Science and Engineering at Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai. My research interests include Artificial Intelligence, Deep learning, Machine Learning, Cyber Security, and Intrusion Detection Systems. I published research papers in reputed journals and conferences. I actively participated in faculty development programs, workshops, and technical seminars. I guided several undergraduate projects and is committed to academic excellence and research innovation.