www.ijarp.com ISSN 2456-9992 Page: 01-13

International Journal Advanced Research Publications

Research Article

Volume: 02
Issue: 04

REAL-TIME AUTOMATED PACKAGING VERIFICATION SYSTEM (RT-APVS) FOR E-COMMERCE LOGISTICS

***Ashutosh, Gagan, Pratham, Jatan / R.D. Engineering College**

Department of Information Technology, Dr. A.P.J. Abdul Kalam Technical University, Uttar Pradesh, India.

Article Received: 19 March 2026, Article Revised: 09 April 2026, Published on: 29 April 2026

***Corresponding Author: Ashutosh**
Department of Information Technology, Dr. A.P.J. Abdul Kalam Technical University, Uttar Pradesh, India.
DOI: https://doi-doi.org/101555/ijarp.5288

ABSTRACT

The accelerated growth of e-commerce has intensified the need for dependable logistical verification systems capable of addressing the increasing financial impact of return fraud. Although conventional interactive video solutions have proven valuable in product marketing and sales contexts, they are not designed to support the operational speed and scale demanded by modern logistics networks. Their dependence on manual configuration and offline video processing limits their practicality in high-volume environments.

To address these shortcomings, this study introduces the Real-Time Automated Packaging Verification System (RT-APVS). The proposed framework employs barcode-triggered mechanisms to embed relevant metadata directly into live video streams at the moment of packaging. By eliminating the need for post-production editing or manual record compilation, RT-APVS streamlines the generation of verifiable digital evidence and strengthens end-to-end traceability across the supply chain. Experimental evaluation indicates that the system reduces dispute resolution time by approximately 90 percent when compared with traditional manual logging approaches.

KEYWORDS: *Supply Chain Security, Dynamic Watermarking, Computer Vision, PHP, MySQL, Logistics, Video Evidence.*

1. INTRODUCTION

1.1 The Digital Transformation of Commerce

The emergence of the Internet has reshaped modern society in ways that are both profound and enduring. Communication, employment, and commercial exchange have increasingly been redefined within an interconnected digital environment. Over the past two decades, the rapid growth in Internet users has fueled an unprecedented expansion of e-commerce, fostering a marketplace that operates beyond conventional physical boundaries. This development represents more than a simple migration from physical stores to online platforms; it reflects a fundamental reconfiguration of the buyer–seller relationship.

As observed by Balfanz et al. [1], organizations have actively adopted the World Wide Web as a strategic channel for product placement, enabling them to engage customers irrespective of geographic constraints. In effect, commerce has become location-independent. A small retailer in Germany can now transact with a consumer in Japan as seamlessly as with a nearby customer. Yet, while this borderless model enhances accessibility and market reach, it also introduces new vulnerabilities.

Unlike traditional transactions—where in-store inspection, payment, and exchange occur simultaneously —online purchases separate these events in time. Payment is often completed well before the goods are physically delivered. This temporal disconnect creates a structural gap in trust, exposing both buyers and sellers to potential exploitation during the interim period.

1.2 The Logistical Complexity Crisis

The rapid expansion of global e-commerce has been accompanied by a parallel increase in logistical complexity. To satisfy rising consumer demand, contemporary supply chains operate through highly coordinated, multi-layered networks. A single shipment may pass through several distinct actors—manufacturers, wholesalers, third-party logistics providers (3PLs), distribution hubs, and last-mile delivery services— before it ultimately reaches the end customer. While these systems are engineered for speed and efficiency, their structural intricacy introduces new operational challenges.

One particularly overlooked dimension within this network is the verification of product fulfillment after the point of sale. Historically, research and investment have concentrated heavily on customer acquisition strategies, including digital marketing, search engine optimization, and interactive sales technologies. In contrast, comparatively limited attention has been directed toward mechanisms that safeguard customer retention through effective dispute resolution and post- delivery validation. As shipment volumes continue to rise, this imbalance becomes increasingly consequential.

1.3 The Multi-Billion Dollar Problem: Return Fraud

Within today's e-commerce environment, return fraud has evolved into a substantial financial burden for retailers and logistics providers alike. Among its various forms, the so-called "empty box" claim—where a customer alleges that an item was missing or damaged upon delivery—has emerged as a particularly costly and pervasive issue. Such claims exploit the structural vulnerability created by the temporal separation between payment and delivery, widening the existing gap of trust between buyer and seller.

Addressing this challenge requires more than conventional shipment tracking systems. A tracking number confirms transit, but it does not verify the condition or contents of a package at the moment it was dispatched. What sellers increasingly require is verifiable, tamper-resistant documentation of the packing process itself—evidence that can accurately attest to the package's contents at the precise point of sealing.

Without such proof, resolving disputes becomes both time-consuming and financially risky.

1.4 The Limitations of Legacy Video-Based Solutions

Several existing video-based approaches have sought primary limitation lies in their dependence on manual authoring tools, such as MOVieEditor, for object tracking and content preparation. As noted in prior research, manual tracking is both insufficient and time-intensive. In a distribution center processing thousands of packages each day, this requirement renders the system impractical.

In addition, these legacy solutions rely on offline processing models. Video content must be edited and compiled into interactive formats before deployment, introducing delays that are incompatible with the need for immediate evidentiary access in logistics operations. Real-time validation is essential when disputes arise, and post-processed footage cannot meet this requirement effectively.

Finally, the user interface paradigm of such systems is misaligned with warehouse workflows. MOVieGoer, for instance, was designed for end-user interaction through mouse-based inputs. In contrast, a packaging verification system must integrate seamlessly with industrial tools—such as barcode scanners—allowing packers to trigger documentation processes without disrupting the pace or structure of manual operations.

1.5 The Proposed Solution: RT-APVS

To address the shortcomings identified in existing systems, this study presents the Real-Time Automated Packaging Verification System (RT-APVS). The proposed framework is

purpose-built for high-volume logistics environments, where speed, accuracy, and evidentiary reliability are essential. Rather than adapting marketing-oriented video architectures, RT-PVS is engineered specifically to support automated documentation at the point of packaging.
To address verification challenges by adapting technologies originally developed for digital marketing. Many of these systems are grounded in architectural models similar to that of MOVieGoer [1], which introduced an innovative framework for integrating interactive video with e-commerce platforms. Through a client–server structure, video elements could be hyperlinked, enabling users to interact directly with products embedded within video content.

Although such systems represented a significant advancement in online retail engagement—for example, allowing viewers to select and purchase items directly from a video—they were not designed with logistical security in mind. Their underlying assumptions and workflows are incompatible with the operational demands of high-volume warehouse environments.

The system builds upon established web technologies, including PHP and MySQL, which have been widely recognized in prior literature as stable and scalable foundations for server-side applications [2][3]. However, the operational model is fundamentally restructured to enable continuous, real-time generation of verifiable evidence. In contrast to interactive sales videos—typically designed for passive viewing and limited user engagement—RT-APVS functions as a data-intensive platform. It actively embeds transactional and packaging metadata directly into the live video stream, thereby transforming the recording into an authenticated evidentiary asset rather than a promotional medium.

2. ARCHITECTURAL EVOLUTION

2.1 The State of Interactive Video in E-Commerce

Hyperlinked or interactive video has been widely examined as a tool for enhancing digital product presentation. Early contributions in this area, notably the MOVieGoer system developed by the Computer Graphics Center, illustrated that video could function as an interactive interface rather than merely a passive viewing medium. This shift redefined the role of video within online retail environments.

A central contribution of MOVieGoer was the introduction of “selectable objects” embedded within video sequences. By enabling users to interact with designated regions of a video frame, the system allowed
retrieval of information, supplementary product including pricing, material specifications, and stock

availability. In effect, video became a navigable environment linked directly to commercial data.

From a technical perspective, this functionality was supported by a structured architectural framework. First, video content was maintained separately from object definitions—typically stored as XML files or database records—to preserve flexibility and facilitate customization. Second, a Database Management System (DBMS), such as MySQL [3], was employed to organize and manage extensive datasets required to associate individual video frames with corresponding product metadata. Finally, dynamic web interfaces generated using server-side scripting technologies like PHP [2], which processed user interactions and retrieved relevant data from the backend in real time. Together, these components enabled a scalable and responsive interactive video ecosystem.

[Image placeholder for Traditional Architecture]
Figure 1. Traditional Interactive Video Architecture (MOVieGoer). The diagram illustrates the structural separation between video content and associated metadata streams. While this modular design enables flexibility and customization, it also introduces distinct data pathways that may increase exposure to security vulnerabilities.

2.2 The "Offline" Bottleneck in Logistics

The architectural principle of data separation offers clear advantages in contexts where information—such as pricing or promotional content—requires frequent modification. However, this same flexibility becomes problematic when applied to evidentiary material that demands immutability and protection. As noted in prior research [1], objects within digital video sequences typically change position over time, necessitating tracking mechanisms to maintain accurate object associations.

Although systems such as MOVieGoer incorporate semi-automated techniques, including key-frame interpolation, as well as automated tracking algorithms like Camshift, these processes are generally conducted after the recording phase within a dedicated editing workflow. Such reliance on post-production processing introduces latency and procedural complexity that are incompatible with the immediacy required in logistics verification.

This limitation becomes particularly significant in dispute scenarios. When a claim arises, sellers must demonstrate precisely what was packed at a specific moment. Any system that depends on subsequent editing or manual designation of "sensitive regions" [5] inherently creates a temporal gap during which data could be altered. If the association between the recorded video and its corresponding Order ID is maintained in a separate, modifiable file—as is the case in traditional architectures—the integrity of that linkage may be compromised. Moreover, earlier studies acknowledge that reliable object tracking remains a persistent Tfeachtonriscal challenge. Collectively, these underscore the inadequacy of offline video-processing models for secure, high-velocity logistics environments.

2.3 The Architectural Pivot

The RT-APVS framework departs from conventional object-tracking paradigms by removing the need for dynamic tracking altogether. In a typical warehouse packing station, the operational environment is spatially stable: the camera remains fixed, and packages are consistently positioned within a predefined area on the packing table. This predictable configuration renders complex object-detection and tracking algorithms unnecessary.

Rather than attempting to identify and follow moving objects within the frame, the system focuses on validating the transactional event occurring within a controlled visual zone. The designated packing area functions as a constant "sensitive region," eliminating variability in spatial reference. Metadata associated with the transaction is embedded directly into the video stream during the encoding process, ensuring that contextual information and visua evidence are integrated at the point of capture. As a result, the recorded file binds transactional data and imagery into a unified, tamper-resistant artifact, thereby strengthening evidentiary integrity.

The RT-APVS framework constitutes a substantive redesign of the interactive video architectures discussed in earlier research. Although it retains the foundational WAMP stack (Windows, Apache, MySQL, PHP) The Video Capture Unit (VCU) consists of a fixed-position, wide-angle camera capable of recording at 1080p resolution and 60 frames per second. Mounted vertically above the packing station, the camera maintains a stable and consistent field of view. This static arrangement removes the need for dynamic object tracking, as the packing zone remains spatially constant throughout operations.

Recognized as a reliable and widely adopted User interaction is facilitated through a Trigger Unit technological standard by Balfanz et al. [1], its (TU), which replaces conventional mouse-based input operational logic differs markedly. Rather than with an industrial barcode or QR code scanner operating preserving the conventional structure, the system restructures the direction and function of data exchange. In Human Interface Device (HID) mode. This configuration allows packers to initiate system actions In traditional implementations such as MOVieGoer, the architecture follows a "pull" model: user interaction initiates a request, and relevant information is retrieved from the backend in response. RT-APVS, by contrast, operates on a "push" paradigm. Metadata is not requested by a user but is automatically embedded into the live video stream when specific physical events occur within the logistics workflow—for example, when a barcode is scanned during the packing process. In this way, evidentiary data is generated proactively and synchronized with real-time operational activity.

[Image placeholder for RT-APVS Architecture]
Figure 2. Proposed RT-APVS Architecture. The diagram illustrates a unidirectional data flow in which information originates at the physical packing station and is transmitted automatically to cloud-based storage, ensuring immediate and continuous synchronization between operational events and digital records.

3. SYSTEM IMPLEMENTATION

3.1 Hardware Layer: The Edge-IoT Configuration

In contrast to the dual-computer architecture—comprising separate server and client systems—employed in earlier interactive video implementations such as MOVieGoer, the proposed framework adopts a consolidated Edge-IoT configuration. This architectural decision reduces transmission latency, lowers hardware expenditure, and within warehouse environments. streamlines deployment By processing and capturing data at the edge of the network, the system enhances responsiveness while maintaining operational efficiency.

through without standard scanning procedures interrupting their workflow or relying on additional input devices such as keyboards or mice. To ensure immediate operational feedback, the system incorporates a Feedback Unit (FU) in the form of a compact LCD display positioned within the packer's line of sight. The display provides real-time status indicators —such as “System Ready” notifications—thereby reinforcing usability and reducing procedural uncertainty during high-volume packaging tasks.

3.2 Software Layer: The Modified WAMP Stack

The software architecture of RT-APVS builds upon the established technologies referenced in prior research while introducing modifications aimed at strengthening system security and operational integrity. Although the foundational components remain consistent with widely adopted web development standards, their functional roles are adapted to meet the evidentiary requirements of logistics environments.

PHP (Hypertext Preprocessor), identified in earlier literature as a prevalent and reliable server-side scripting language [2], is employed to manage HTTP POST requests generated during packing events. Its execution on the server side—where source code remains inaccessible to end users—adds a structural layer of protection against unauthorized manipulation. Within RT-APVS, PHP coordinates transactional data flow between the edge device and backend services.

To organize and store metadata at scale, the system utilizes a MySQL database, recognized for its performance and stability in managing large datasets [3]. The database schema is designed to establish explicit relational links among key variables, including Order_ID, Video_File_Path, Timestamp, and Packer_ID. This structured mapping ensures traceability and facilitates rapid retrieval during dispute resolution.

A notable divergence from earlier architectures, such as the Java applet-based player used in MOVieGoer, lies in the adoption of a Python-based processing layer. RT-APVS deploys a local Python script integrated with the OpenCV library to capture live video streams and embed metadata directly into each frame during recording. By performing frame-level text overlay in real time, the system transforms the video file into an inseparable record of both visual and transactional data, thereby enhancing evidentiary robustness.

3.3 The Core Algorithm: Dynamic Overlay

Whereas the MOVieGoer architecture processes user-generated interaction data—such as mouse coordinates—to retrieve information from a database, the RT-APVS framework operates on an alternative logic. Instead of querying data in response to on-screen interaction, the system captures trigger data in the form of scanned barcode strings and immediately embeds the corresponding metadata into the video stream.

This shift transforms the role of input events within the system. Barcode scans do not initiate a request–response cycle; rather, they activate a write operation that integrates transactional information directly into the recorded footage. The algorithm responsible for this dynamic overlay executes locally on the edge device, thereby minimizing network dependency and eliminating manual logging workflow and the proposed automated framework.

Dataset A (Manual/Legacy Approach):

In the first group, operators followed a procedure analogous to the MOVieGoer-style authoring model. Each packaging session was recorded manually, after which the operator stopped the recording and subsequently used a separate software interface to annotate or "tag" the video with the corresponding Order ID. This methodology reflects authoring processes where an evidentiary linkage is established only after the recording phase has concluded.

Dataset B (RT-APVS/Automated Approach):

In the second group, operators employed the RT-APVS system. Recording was initiated and terminated exclusively through barcode-based triggers: scanning the order barcode commenced the session, and scanning the shipping label concluded it. No direct interaction with the workstation was required. This configuration ensured that metadata integration occurred automatically during capture, eliminating the need for post-recording annotation.

4.2 Mathematical Model for Efficiency

To quantify operational efficiency, the cycle time for processing a single package is defined as:

$$Ct = T_{pack} + T_{overhead}$$

processing delays. By confining computation to the point

Here, T_{pack}

represents the time required to physically of capture, the system ensures near-zero latency and preserves synchronization between physical packing events and their digital documentation.

4. METHODOLOGY

To assess the effectiveness of RT-APVS, a controlled experimental study was undertaken. The evaluation compared the proposed automated framework with a Manual Logging Method, which reflects the manual authoring approaches described in earlier literature. By examining both systems under comparable operational conditions, the study aimed to determine differences in place the items into the package. This component remains constant across both experimental groups, as the packing procedure itself does not vary. In contrast, *Toverhead* denotes the additional time associated with system interaction.

Within the Manual Logging Method (Dataset A), this overhead arises primarily from the post-recording tasks required to link video evidence with the corresponding Order ID. Such steps include manual stopping of recordings and subsequent tagging or annotation through a separate interface. Consistent with prior observations regarding the time-intensive nature of manual tracking, efficiency, performance. reliability, and dispute-resolution
this interaction introduces measurable delays that directly affect overall cycle time.

4.1 Dataset Generation

For experimental evaluation, a proprietary dataset titled Pack-Logistics-5k was developed, comprising 5,000 simulated packaging sequences. The dataset was structured to facilitate a comparative analysis between a

5. RESULT

5.1 Latency and Efficiency Analysis

The primary bottleneck identified in the legacy MOVieGoer architecture was the reliance on manual or

6.1 Addressing the Tracking Constraint

Prior research has acknowledged that object tracking constitutes one of the principal challenges in authoring semi-automatic processes, where "Tracking an object interactive video systems [1]. Conventional manually in each frame of a sequence is insufficient and implementations relied either on computationally time consuming". Our data confirms this observation in a logistical context.

6. DISCUSSION

Dataset A (Manual): The average overhead per package was recorded at 45.2 seconds. This included the time required to stop the recording, manually rename the file to match the invoice number, and initiate the upload. In high-volume environments, this latency is prohibitive.

Dataset B (RT-APVS): The average overhead was recorded at 0.2 seconds, corresponding strictly to the latency of the barcode scanner and the PHP script's "handshake" with the database.

5.2 Database Performance and Retrieval

The performance of the backend infrastructure was closely tied to the efficiency of the MySQL database server [3], which managed video file paths and associated transactional metadata. Even after populating the database table with 5,000 indexed records, system responsiveness remained consistently high.

Empirical testing indicated that the PHP-based retrieval script located and returned the appropriate video path in an average time of 0.012 seconds, demonstrating minimal latency under the evaluated workload. These results suggest that the database architecture scales effectively for moderate dataset sizes without compromising retrieval speed.

In addition to query performance, the system's capacity for dynamic data updates was examined. As highlighted in earlier literature [1], associated information can be modified without altering the underlying video content. To validate this functionality, the status field of 100 selected orders was changed from i““nStheripfapceed” to “Disputed.” The retrieval reflected these updates immediately, flagging the relevant records without necessitating any modification or reprocessing of the stored video files. This outcome confirms the separation between evidentiary media and mutable transactional metadata at the database layer, thereby preserving both operational flexibility and data integrity.

intensive algorithms, such as Camshift, or on labor-intensive manual key-framing techniques to define and maintain "selectable objects" across video sequences.

Both approaches introduce complexity and limit scalability, particularly in high-throughput environments. Within the logistics context targeted by RT-APVS, this constraint is effectively circumvented through architectural simplification. The camera perspective is fixed, and the packing zone remains spatially stable tdhyrnoaumghicoaultly operations. Rather than identifying and tracking the parcel within the frame, the system associates evidentiary recording with a clearly defined temporal window—specifically, the interval between the initiation and termination barcode scans.

By anchoring verification to time-based triggers rather than spatial detection, RT-APVS achieves complete temporal alignment between the transaction event and the recorded footage. This approach eliminates the need for computationally expensive computer vision tracking algorithms while maintaining evidentiary reliability. The system does not require precise localization of the parcel within the frame; it is sufficient to confirm that the packaging activity occurs within the predefined visual field during the transaction interval.

6.2 Architectural Security and System Integrity

The adoption of a WAMP/LAMP-based architecture acnodtreibvuidteenstiasryig innitfeigcarnittyl.yAs thoighsliygshtetend inspecriuomritlyiteratue [1], PHP operates as a server-side scripting language in which source code remains concealed from the client; it is executed on the server and not transmitted to the user's browser. This characteristic is particularly important in the context of an evidence-generation platform.

In RT-APVS, the core logic governing metadata validation, recording triggers, and database updates resides entirely within server-side PHP scripts. Unlike client-side technologies such as JavaScript—where executable logic can be inspected or manipulated—the server-centric design restricts direct access to operational code. Consequently, potential attackers are unable to examine or alter the mechanisms responsible for initiating or associating video evidence, thereby reducing the likelihood of circumvention or exploitation.

6.3 Scalability

The system architecture also supports efficient scalability across distributed packing stations. Because aarpeplication logic and content management centralized on a dedicated server rather than dispersed across heterogeneous client devices, updates can be deployed uniformly and instantaneously. Modifications to recording protocols or compliance requirements propagate automatically to all connected stations without requiring local reconfiguration.

This centralized control model parallels the controlled update mechanisms discussed in earlier interactive video systems, while adapting them to the evidentiary domain.

As a result, RT-APVS ensures that all packaging verification processes operate under consistent and current standards, even as operational scale increases.

VII. CONCLUSION

The application of video technologies within e-commerce has historically concentrated on front-end engagement, emphasizing interactive catalogs and hyperlinked product displays to enrich the consumer experience. This study illustrates that the same foundational architectural concepts—particularly client–server data separation and the deployment of robust database management systems such as MySQL—can be strategically reoriented to address a fundamentally different concern: logistical verification at the back end of the supply chain.

The proposed RT-APVS framework resolves the operational constraints associated with offline authoring and manual object tracking that characterize earlier systems. By replacing user-driven interaction with barcode-triggered automation and restructuring the data flow to embed metadata directly into the live video stream, the system transforms evidentiary recording into a real-time, integrated process. This approach not only reduces procedural latency but also strengthens resistance to tampering.

Empirical evaluation indicates that adapting a proven PHP/MySQL infrastructure for packaging verification yields reliable and accurate outcomes. More importantly, it redefines the role of packaging video within e-commerce operations. Rather than serving as a passive archival record, the video becomes a structured, searchable, and verifiable asset capable of supporting dispute resolution and mitigating return fraud at scale.

REFERENCES

1. D. Balfanz, M. Finke, C. Jung, and R. Wichert, "An interactive Video System supporting E-Commerce Product Placement," Computer Graphics Center, ZGDV, Darmstadt, Germany.
2. PHP Group, "PHP Scripting Language," http://www.php.net.
3. MySQL AB, "MySQL," http://www.mysql.com.
4. Apache Software Foundation, "Apache Server," http://www.apache.org.
5. V. Burrill, T. Kirste, and J. Weiss, "Time-varying Sensitive Regions in dynamic multimedia objects," Information and Software, 36(4), 1994.
6. C. Freytag, C. Kumpf, and L. Neumann, "Utilization of device and user profiles for mobile WWW-Access," MoMuC'98, Berlin, 1998.
7. N. Sawhney, D. Balcom, and I. Smith, "HyperCafe: narrative and aesthetic properties of hypervideo," Proceedings of the 7th ACM conference on HYPERTEXT, 1996.