International Journal of Management Science and Information Technology
IJMSIT

E-ISSN: 2774-5694 P-ISSN: 2776-7388 OPEN ACCESS

Volume 6 (2), July-December 2026, 3322-3337
DOI: https://doi.org/10.35870/ijmsit.v6i2.9017

A Comprehensive AI-Driven Cyber-Physical Security Framework for ATM Booths with Forensic Readiness: A Bangladesh Perspective

Rathindra Nath Mondal [1*]
[1*] IT Auditor and independent researcher, University of Dhaka, Dhaka 1000, Bangladesh

Email: rathindra.live@gmail.com [1*]

Article history:
Received October 17, 2026
Revised November 15, 2026
Accepted November 24, 2026

Abstract

Automated Teller Machines (ATMs) constitute essential cyber-physical banking infrastructure, particularly in developing economies such as Bangladesh. Nevertheless, they remain vulnerable to physical manipulation, card skimming, social engineering, and sophisticated cyber-physical threats, while existing security solutions rely on reactive CCTV monitoring and rule-based alarms with limited intelligence and forensic capability. This study introduces an intelligent security framework combining IoT sensors, computer vision, machine learning-based behavioural analysis, microcontroller-driven actuators, cloud-based notifications, and forensic-ready data logging. The proposed approach facilitates proactive threat identification, automated incident response, reduced false alarms, automated responses (e.g., door locking, fire detection), SMS/phone alerts, and reliable evidence preservation. Experimental evaluation achieved 94.3% detection accuracy, 2.1-second response latency, 99.1% system uptime, and 0.7-second cloud communication delay. Furthermore, ANOVA-based statistical assessment at a 95% confidence level demonstrated significant performance improvements. Aligned with ISO/IEC 27043:2015, the framework offers a comprehensive approach to strengthening ATM security, operational resilience, and post-incident forensic investigations.

Keywords:
ATM Booth Security; Cyber-Physical Systems; Artificial Intelligence; Forensic Readiness; IoT; Bangladesh.

1. INTRODUCTION

ATMs support 24/7 banking and financial inclusion in Bangladesh, but their growing deployment increases security threats, including card skimming, vandalism, unauthorized access, and financial fraud.

Traditional ATM security measures, including CCTV surveillance, access controls, and manual monitoring, offer limited preventive capabilities and forensic preparedness, often identifying threats after incidents occur (Reddy, 2023). The increasing sophistication of financial fraud and cyber-physical attacks demands adaptive security solutions (Pise et al., 2020). Leveraging AI, IoT, and computer vision, this study proposes an integrated AI-driven cyber-physical security framework for ATM booths in Bangladesh (Thopate et al., 2023). The main contributions of this research are:

a. Development of a comprehensive cyber-physical security framework integrating AI, IoT, and forensic readiness for ATM booths.
b. Formulation of a Bangladesh-specific ATM booth threat model.
c. Design of AI-based behavioural analytics and sensor fusion mechanisms.
d. Evaluation of the proposed framework through simulated ATM booth security scenarios.

1.1. Literature Review

ATM security research has evolved from CCTV and alarm-based physical protection (Thiel, 2000) to biometric authentication using fingerprint and facial recognition (Jacintha et al., 2017). However, these approaches inadequately address real-time booth-level threats. Nurdi (2019) demonstrated that combining

machine learning classifiers (Random Forest, LightGBM, CatBoost) with CNN and LSTM improves anomaly detection and reduces false alarms, highlighting the potential of integrated intelligent security mechanisms.

Machine learning has improved transaction-level fraud detection (Gavaskar et al., 2022), while computer vision enables suspicious activity recognition, including unauthorized access and prolonged presence (Porretti et al., 2016). Deep learning and IoT-based systems further demonstrate real-time anomaly detection through video analytics, facial recognition, and motion sensors, triggering automated alerts and interventions (Deshpande et al., 2018; Arroyo et al., 2015; Venugopal, 2019). Additionally, Abiew et al. (2020) highlighted the effectiveness of deep learning, ensemble, and hybrid models in detecting complex digital banking fraud patterns.

Forensic readiness involves proactively collecting and preserving digital evidence for investigations and legal proceedings (Ramyasri & Mahalakshmi, 2020), yet its application to ATM security remains limited (Kumar et al., 2024). Similarly, existing PIN-inference research primarily examines vulnerabilities rather than real-time AI-driven detection and prevention. The lack of automated intervention and cyber-physical integration highlights a significant research gap (Perera et al., 2020), which this study addresses through an integrated AI-driven ATM security framework incorporating forensic readiness.

1.2. Types of ATM Booth Frauds Observed in Practice during IT Audit

Table 1 summarizes ATM fraud types, impacts, and detection challenges observed worldwide and in Bangladesh. This taxonomy reflects evolving cyber–physical threat patterns and motivates the need for intelligent, forensic-ready security frameworks (Ghose et al., 2025).

Table 1. Common types of ATM booth frauds

Sl.	Fraud Type	Description / Modus Operandi	Impact on Customers & Banks	Detection Challenges
1	Card Skimming	Installation of skimming devices on card slots to capture card data and PINs	Financial loss, identity theft, loss of customer trust	Devices are small and hard to detect; often installed during low-traffic hours
2	Shoulder Surfing	Criminals observe customers entering PINs, often in crowded or poorly designed booths	Account compromise, unauthorized withdrawals	Difficult to detect using transaction logs alone; requires behavioural monitoring
3	Cash Trapping	Fake cash outlets trap dispensed cash, later retrieved by attackers	Direct financial loss to customers	Appears as successful transaction in system logs
4	Card Trapping	Devices retain customer cards to enable later misuse	Card misuse, service disruption	ATM reports card capture as normal error
5	ATM Malware Attacks	Malware installed to control ATM cash dispenser (e.g., jackpotting)	Large-scale cash loss, reputational damage	Requires correlation between cyber and physical events
6	Forced Withdrawal / Coercion	Victims are forced to withdraw cash under threat	Severe personal and financial risk	Victim behaviour may appear normal to systems
7	Vandalism & Physical Tampering	Damage to ATM components or booth infrastructure	Service downtime, repair cost	Often detected only after damage occurs
8	Insider-Assisted Fraud	Bank or service personnel collude with criminals	High financial and reputational risk	Legitimate access masks malicious intent
9	SIM Swap–Linked ATM Fraud	SIM swap enables OTP interception for ATM transactions	Unauthorized withdrawals	Telecom–bank coordination required
10	Fake Assistance Fraud	Attackers pose as helpers to deceive ATM users	Account compromise, data theft	Social engineering difficult to detect technically
11	Explosive / Mechanical Attacks	Use of gas or mechanical force to break ATM safes	Massive financial loss, public safety threat	Rare but highly destructive; rapid detection required
12	Booth Loitering & Surveillance	Criminals monitor customer behaviour before attack	Increased likelihood of targeted fraud	Requires long-term behaviour analysis

1.3. Comparative Analysis of Existing ATM Security Approaches

Table 2 compares major existing ATM security approaches reported in prior research and industry deployments, highlighting strengths, limitations, and motivating an integrated AI-driven, forensic-ready framework.

Table 2. Existing ATM security approaches

Sl.	Security Approach	Key Highlights	Strengths	Limitations
1	Sensor-Based Security Systems	Buzzer alerts, SMS notifications, location status tracking, automatic door lock/close activation	Effective for basic physical intrusion detection; low deployment cost; immediate local alerts	Limited to physical threats only; lacks intelligence to distinguish false alarms; no behavioural analysis or forensic evidence generation
2	Transaction-Level Security Systems	Multi-factor authentication, PIN + OTP verification, transaction validation rules	Strong protection against financial fraud and unauthorized transactions; effective for account-level security	Does not address booth-level physical threats; cannot detect pre-transaction crimes such as skimming or coercion; no situational awareness
3	Machine Learning–Based Security Systems	Face recognition–based user authentication, anomaly detection in transaction patterns	Improves user verification accuracy; adaptive learning capabilities; reduces identity fraud	Computational overhead may introduce transaction delays; often ignores physical and environmental threats; limited forensic readiness
4	Camera-Based Surveillance Systems	Detection of masked individuals, weapons, suspicious movement using CCTV	Visual evidence availability; deterrence against visible crimes; useful for post-incident review	Predominantly reactive; delayed response due to manual monitoring; fragmented evidence and weak chain-of-custody
5	Access Control & Smart Lock Systems	Time-based locking, RFID-enabled access, remote lock/unlock	Prevents unauthorized booth entry; simple automation	No intelligence on human behaviour; ineffective against insider misuse; no evidence correlation
6	Panic Button & Manual Alarm Systems	Emergency button activation by users or guards	Simple and low-cost; immediate alert generation	Depends on human intervention; ineffective during sudden attacks; no automated analysis
7	GPS & ATM Tracking Systems	ATM location tracking, movement detection	Useful against ATM theft or relocation; supports asset recovery	Ineffective for in-booth crimes; no real-time behaviour detection
8	Rule-Based Video Analytics	Predefined rules for motion, intrusion, or crowd detection	Easy to deploy; low computational complexity	Rigid rules; high false alarms; poor adaptability to evolving attack patterns
9	Isolated IoT Monitoring Systems	Independent monitoring of door, vibration, smoke sensors	Early detection of mechanical tampering; real-time alerts	Lack of cross-validation with vision; vulnerable to false triggers
10	Proposed AI-Driven Cyber–Physical & Forensic-Ready Framework	CNN-based behaviour analysis, LSTM anomaly detection, sensor fusion, forensic logging	Proactive threat detection; low false alarms; audit-ready evidence; regulatory alignment	Higher initial deployment cost; requires skilled maintenance

1.4. Timeline of Incidents of Bangladeshi Banking Sector

Table 3. Timeline of Incidents of Bangladeshi Banking Sector

Year	Bank / Institution	Threat Agent / Mechanism	Incident	Impact
2012	United Commercial Bank (UCB)	Skimmer, Fraudster	Card Cloning, ATM forgery, Buying	BDT 100 million

			product from shopping malls	
2013	Mutual Trust Bank (MTB)	IT officials and their associates	ATM forgery, Employee fraud	BDT 20 million
	Sonali Bank Ltd.	Hackers	Malware and Keylogger Malware	$0.25 million
	Bangladesh Bank Heist	Hackers	Keylogger, manipulating the SWIFT system	$81 million
2016	City Bank, Eastern Bank, United Commercial Bank (ATMs – 40 cards cloned)	Skimmer, Miscreants	Card Cloning, Exploiting weaknesses in credit and debit card system, ATMs, points of sale(POS) and banks, phishing	BDT 2.5 million
	Foreign origin; Non-branded Malaysian financial institution (Premier Bank)	Fraudster	Card Cloning, ATM forgery	BDT 3.4 million
	Foreign origin; Cloned cards from Riyadh Bank (Prime Bank)	Skimmer, Fraudster	Card Cloning, ATM forgery	BDT .5 million
	DBBL (ATM)	Hackers	Malware, ATM robberies	BDT 1.2 million
2019	Dutch Bangla Bank Ltd.	Hackers	Malware	$3 million
	Bangladesh Bank	Hackers	Ransomware, Malware	Info breach
2021	Bangladesh Bank	Hackers	Malware	Info breach
	DBBL (Dutch Bangla Bank Limited) – ATM Fraud	Bank Employee, Fraudsters	ATM card theft, Charging Electric Journal	BDT .25 million
2022	Many banks (survey)	Malware, server intrusion, data theft	Widespread cyber-attacks on banking systems	Not quantified individually (survey) (en.bd-pratidin.com)
2023	Banking sector	Malware, server intrusion, data theft	Continued cyber-attack prevalence	Broad systemic fraud/exposure (en.bd-pratidin.com)
2024	Nagad / Bangladesh Bank oversight	Financial irregularities / unauthorized e-money	Nagad Ltd fraud case (filed early 2025)	~ Tk 645 crore allegation (The Daily Star)
2025	Bangladesh Bank NSC system	Internal system breach / password misuse	National Savings Certificate embezzlement	~ Tk 25 lakh embezzled, ~Tk 50 lakh attempt blocked (The Business Standard)
2025 (other)	Bangladesh Bank reports	Multiple unidentified mechanisms	Surge in suspicious activities	56% increase in reports (The Business Standard)

1.5. Research Questions and Hypotheses

This study is guided by the following research questions (RQs) and hypotheses (Hs).

RQ1: Can AI-driven behavioural analytics significantly improve ATM booth threat detection accuracy compared to traditional CCTV-based monitoring?

H1: AI-based behavioural models achieve statistically higher detection accuracy and lower false alarm rates than rule-based CCTV monitoring.

RQ2: Does cyber-physical sensor fusion enhance early detection of ATM booth security incidents?

H2: Multi-sensor fusion (IoT + vision) reduces detection latency and false positives compared to single-source monitoring.

RQ3: How does forensic-ready logging improve post-incident investigation and audit effectiveness?

H3: Forensic-ready architectures provide higher evidence completeness and reconstruction accuracy than conventional security logging.

1.6. Threat Model and Bangladesh Context

ATM booth security threats in Bangladesh can be broadly categorized into physical, cyber, and operational threats.

1.6.1. Physical Threats

Physical threats include forced entry, vandalism, tampering, skimming, and explosive attacks, worsened by limited surveillance and delayed response.

1.6.2. Cyber and Fraud-Related Threats

Cyber threats include skimming, data interception, malware, shoulder surfing, and insider-assisted attacks, complicating prevention.

1.6.3. Operational Challenges

Bangladesh's ATM security faces inadequate monitoring, manual dependency, and limited forensic evidence, demanding intelligent integration.

1.7. Proposed AI-Driven Cyber-Physical Security Framework

The proposed framework adopts a layered architecture combining sensing, perception, intelligence, and forensic response components.

1.7.1. Model Configuration and Hyperparameters

1.7.1.1. CNN Architecture

ResNet-50 extracts robust 2048-dimensional visual features from 224×224 ATM video frames using its global average pooling layer.

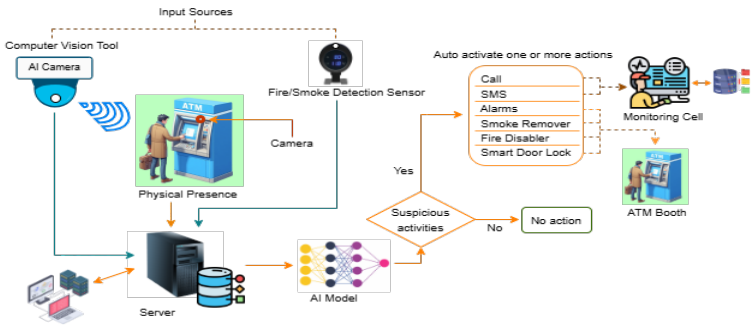

Figure 1. AI-Driven Cyber-Physical Security Framework

LSTM temporal modelling, to capture sequential behavioural patterns across video frames, the extracted CNN features are passed to a LSTM network. The LSTM configuration is defined as table 4.

Table 4. The LSTM configuration

Parameter	Value
LSTM layers	2
Hidden units per layer	128
Sequence length	20 frames
Dropout rate	0.3
Activation	Tanh
Optimizer	Adam
Learning rate	0.001

The LSTM network models temporal dependencies in user movement patterns to detect abnormal behaviours.

Table 5. Model Hyperparameters

Hyperparameter	Value
Batch size	32
Number of epochs	50
Learning rate	0.001
Loss function	Binary Cross-Entropy

Optimizer	Adam
Input frame resolution	224 × 224

These hyperparameters were selected empirically to balance training stability and detection accuracy.

1.7.1.2. Sensing Layer

The sensing layer uses IoT motion, door, vibration, and environmental sensors to monitor ATM conditions continuously.

1.7.1.3. Perception Layer

AI-enabled CCTV and computer vision detect occupants, count individuals, and recognize suspicious behaviours, including loitering and forced entry.

1.7.1.4. Intelligence, Response and Forensic Layer Layer

CNN-based intelligence combines visual and sensor data for real-time threat assessment, triggering alerts, alarms, emergency lockdowns, and forensic evidence logging.

2. RESEARCH METHOD

The methodology in this study follows three main steps: AI-Based Behavioural Analytics, Sensor Fusion, and Forensic Readiness.

2.1. System Overview and Data Flow

The proposed system operates on continuous multi-source data ingestion and real-time inference. Data streams originate from ATM booth CCTV video feeds, ATM machine–embedded cameras, and Physical safety sensors such as fire, smoke, vibration, and door access sensors.

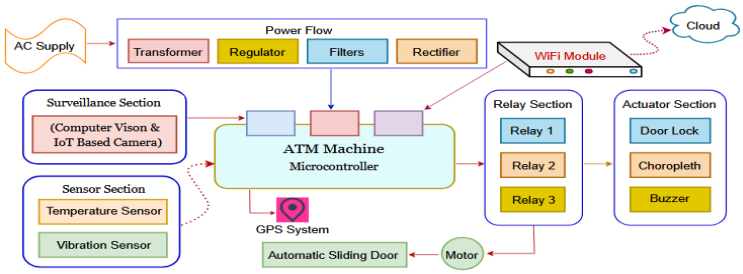

Figure 2. IoT based real-time ATM surveillance and security system diagram

2.2. Real-Time Computer Vision–Based Behavioural Analytics

2.2.1. Video Stream and Dataset Composition

The visual dataset comprises real-time video streams from two sources: External ATM booth CCTV cameras capturing full-scene activity, and ATM machine–embedded cameras capturing frontal facial views during transactions.

Dataset Summary, the experimental dataset includes ATM surveillance videos and synchronized sensor events from simulated environments, segmented into clips reflecting diverse ATM security behaviours.

Table 6. Dataset Summary

Dataset Attribute	Value
Total number of video clips	12,480
Total hours of CCTV footage	38.6 hours
Average clip duration	5–60 seconds
Frame rate	25 fps
Total sensor events recorded	9,350

Table 7. Sensor Event Distribution

Sensor Type	Event Count
Motion detection	3,200
Door open/close events	2,150
Vibration/tampering detection	1,480
Fire/smoke alerts	320
ATM internal camera triggers	2,200

Total	9,350

2.2.2. Face Matching and Behavioural Profiling

Encrypted facial and behavioral profiles enable anomaly detection through identity matching and activity deviations, supporting risk assessment (Mishra, 2023).

2.2.3. Movement and Interaction Analysis

Human pose estimation and trajectory analysis track ATM booth movements, flagging anomalies like excessive motion, repeated entries, or forced gestures, which are fused with facial verification to determine user intent.

2.2.4. Video Stream Acquisition and Pre-processing

ATM CCTV streams are captured, resized, normalized, denoised, and motion-segmented to robustly focus on human regions under varying lighting conditions.

2.2.5. Human Detection and Occupancy Analysis

A deep learning object detection model identifies humans and counts occupants to detect abnormal occupancy, tailgating, or prolonged loitering in ATM booths.

2.2.6. CNN-Based Feature Extraction

Each frame is processed via CNN to extract posture, movement, proximity, and interaction features, forming a temporal behavioural sequence inside the ATM booth.

2.2.7. Temporal Anomaly Detection Using LSTM and Autoencoders

An LSTM autoencoder trained on CNN-extracted features of normal ATM behaviour captures temporal patterns; high reconstruction errors during real-time monitoring indicate anomalies like forced withdrawals, loitering, or coordinated suspicious activities.

2.3. Cyber–Physical Sensor Fusion for Context-Aware Risk Assessment

2.3.1. IoT Sensor Integration

Multiple IoT sensors—door, vibration, motion, and environmental—continuously monitor the physical state of the ATM booths, detecting tampering, forced entry, and abnormal physical events.

2.3.2. Fusion Logic and Risk Scoring

Vision-based anomaly outputs are fused with IoT sensor readings via a weighted risk score, prioritizing high-impact events per forensic IT audit, producing a comprehensive cyber-physical threat assessment for ATM booths.

2.3.2.1. Mathematical Risk Scoring Model for Cyber-Physical Threat Detection

The system classifies ATM events as legitimate, suspicious, or alarming by fusing computer vision analytics with IoT sensor data into a multi-modal risk score. Let the ATM booth monitoring system receive observations from three primary sources:

a. Visual anomaly score from the CNN–LSTM behavioural model.
b. IoT sensor signals (motion, door, vibration, smoke, fire detectors).
c. Customer identity verification and behavioural consistency.

The overall risk score at time t is defined as:

$$\mathcal{R}_t = \alpha\mathcal{V}_t + \beta\mathcal{S}_t + \gamma\mathcal{B}_t$$

Where:
$\mathcal{R}_t$ = overall cyber-physical risk score
$\mathcal{V}_t$ = visual anomaly score extracted from video analytics
$\mathcal{S}_t$ = IoT sensor anomaly score
$\mathcal{B}_t$ = behavioural inconsistency score based on customer identity and movement pattern
α, β, γ = weighting coefficients such that

$$\alpha + \beta + \gamma = 1, \qquad \alpha, \beta, \gamma \geq 0$$

The weights α, β, γ were optimized using grid search on a validation dataset to maximize F1-score while minimizing false alarm rate, subject to $\alpha + \beta + \gamma = 1$. The final values ($\alpha = 0.5, \beta = 0.3, \gamma = 0.2$) emphasize visual anomaly detection while incorporating sensor and behavioural cues in the fusion model.

IoT sensors produce binary or continuous readings representing physical events in the ATM booth. The aggregated sensor anomaly score is computed as:

$$S_t = \sum_{i=1}^{n} w_i s_i$$

Where:
s_i = normalized value from sensor i
w_i = importance weight assigned to each sensor
n = total number of sensors

Example sensors included are motion detector, door access sensor, vibration/tampering detector, fire/smoke sensor, and ATM machine internal camera trigger.

The visual anomaly score is obtained using reconstruction error from the anomaly detection model:

$$V_t = ||X_t + \hat{X}_t||^2$$

Where:
X_t = original feature vector extracted from the CNN
$\hat{X}_t$ = reconstructed feature vector produced by the LSTM autoencoder

Higher reconstruction error indicates abnormal behaviour such as loitering, tailgating, forced entry, and suspicious movement patterns.

The behavioural score evaluates whether the current user behaviour matches historical behaviour patterns associated with the authenticated customer.

$$\mathcal{B}_t = 1 - sim(F_{current}, F_{historical})$$

Where:
$F_{current}$ = current behavioural feature vector
$F_{historical}$ = historical behavioural profile of the user
$sim(\cdot)$ = similarity function (cosine similarity)

A large deviation indicates potential impersonation or suspicious activity.

The final system decision is determined by comparing the computed risk score $\mathcal{R}_t$ with predefined thresholds:

$$Decision = \begin{cases} Legitimate, \mathcal{R}_t < \theta_1 \\ Suspicious, \theta_1 < \mathcal{R}_t < \theta_2 \\ Alarming, \mathcal{R}_t \geq \theta_2 \end{cases}$$

Where:
θ_1 = suspicious activity threshold
θ_2 = critical alarm threshold

When $\mathcal{R}_t \geq \theta_2$, the system automatically triggers:
a. Alarm activation
b. SMS / phone alert to monitoring center
c. ATM booth door lock
d. Forensic event logging

Benefits of the Risk Scoring Model, the proposed sensor-fusion risk model offers several advantages:
a. Reduced false alarms by combining visual and physical signals
b. Faster detection of coordinated cyber-physical attacks
c. Adaptive security response through risk-based thresholds
d. Forensic traceability by logging intermediate scores and sensor events

This mathematical formulation enables real-time threat detection while supporting forensic investigation and regulatory auditing requirements for ATM security systems.

2.4. Automated Response: AI-Driven Door Locking and Alerts

When the computed risk score exceeds a predefined threshold, the system automatically initiates protective actions. These include:

a. Immediate locking of ATM booth doors to prevent further intrusion.
b. Activation of audible alarms.
c. Real-time alerts to bank security operation centres.

AI-driven predictions directly trigger automated door locks, enabling rapid, manual-free responses during critical ATM events.

2.5. Forensic Readiness and Secure Evidence Management

2.5.1. Event Logging and Metadata Capture

Only event-triggered, anonymized video snippets are securely stored for forensic purposes; continuous raw video is not retained.

2.5.2. Evidence Integrity and Chain-of-Custody

Forensic logs use cryptographic hashing and access control, ensuring evidence integrity for audits, compliance, and legal proceedings.

2.5.3. Role of Forensic IT Audit Expertise

Forensic IT audit experience guides indicator selection, threshold setting, and evidence retention, ensuring threat detection produces audit-ready, legally defensible evidence per Bangladesh Bank requirements.

2.6. Forensic Readiness Design and IT Audit Perspective

2.6.1. Forensic Event Taxonomy

The framework logs access violations, abnormal occupancy, forced entry, sensor tampering, and automated responses with metadata.

2.6.2. Evidence Integrity and Chain-of-Custody

All forensic logs are protected using cryptographic hashing, synchronized timestamps, and secure storage to preserve evidence integrity. Chain-of-custody principles ensure that evidence remains admissible for internal investigations and legal proceedings.

2.6.3. Role of Forensic IT Audit Experience

Forensic IT audit principles guide event selection, logging, and evidence retention, strengthening accountability, incident reconstruction, regulatory compliance, and ATM security.

2.6.4. Mapping to Bangladesh Bank Audit Expectations

The framework supports Bangladesh Bank ICT compliance through continuous monitoring, access logging, incident traceability, and audit-ready reporting, reducing institutional audit risks.

2.7. Experimental Setup and Evaluatoin

2.7.1. Experimental Environment

A hybrid ATM testbed evaluated detection accuracy and latency using recorded/live video, simulated transactions, and IoT sensors under realistic deployment conditions.

2.7.2. Dataset Description, Partitioning and Regulatory Alignment

2.7.2.1. Dataset Size and Composition

The dataset includes 12,480 ATM video segments (5–60 seconds, 25 fps) collected over six months from urban and semi-urban ATM locations in Bangladesh, categorized into table 8.

Table 8. Class Distribution

Activity Class	Number of Clips	Percentage
Legitimate Activity or Normal Transactions	11,907	95.42%
Suspicious Activity (PIN exposure / shoulder surfing risk)	150	1.20%
Alarming Events	423	3.38%
Total	12,480	100%

Stratified 10-fold cross-validation was used to mitigate class imbalance and ensure fair performance comparison. The dataset consists of three synchronized components:

a. Video Dataset: ATM booth AI cameras video footage, CCTV videos and ATM camera face videos collected from public surveillance datasets and simulated scenarios, designed using real-world security incident patterns observed during professional IT security incident audit engagements. Video clips range

from 5–60 seconds and include normal usage, suspicious loitering, masked intrusion, forced transactions, and vandalism scenarios.

b. Sensor Dataset: Time-stamped fire, smoke, vibration, and door access sensor signals generated using controlled simulations and historical incident patterns.
c. Behavioural Profile Dataset: Historical customer transaction duration, movement patterns, and facial embeddings.

The dataset was split into training (70%), validation (15%), and testing (15%) sets, ensuring class balance across legitimate, suspicious, and alarming categories.

2.7.2.2. Data Sources and Privacy Compliance

The dataset was obtained through controlled collaboration with banking institutions and simulated experimental setups. All recordings were anonymized by:

a. Blurring facial regions.
b. Masking card numbers and transaction details.
c. Removing personally identifiable information (PII).

The study complies with institutional research ethics and digital surveillance data handling guidelines.

2.7.2.3. Annotation Process

A three-stage annotation protocol was adopted in this study: 1) Independent labelling by two trained forensic IT auditors, 2) Consensus review for conflicting labels (~6.4%), 3) Expert validation for ambiguous cases. Inter-annotator reliability measured using Cohen's Kappa yielded: $\mathcal{K} = 0.87$. Indicating strong agreement and high labelling consistency.

2.7.2.4. Regulatory Alignment with Bangladesh Bank

The dataset and features align with ATM security risks, audit findings, and Bangladesh Bank ICT security guidelines. Specifically, the framework addresses:

a. CCTV monitoring weaknesses.
b. PIN visibility exposure risks.
c. Inadequate booth access control.
d. Delayed fraud detection due to manual review.

The proposed behavioural analytics and multi-sensor fusion support Bangladesh Bank ICT compliance, enhancing fraud detection and ATM risk mitigation.

2.8. Model Training Configuration

Pre-trained CNNs extracted visual features, while LSTM–Autoencoders learned temporal patterns; normalized sensor data, combined loss functions, early stopping, and k-fold validation minimized overfitting.

2.9. Evaluation Metrics

System performance was evaluated using both classification and operational metrics: 1) Accuracy, Precision, Recall, and F1-score for multi-class threat classification, 2) False Alarm Rate (FAR) to measure unnecessary alerts, 3) Detection Latency, defined as time between anomaly occurrence and system response, 3) Area Under ROC Curve (AUC) for binary anomaly detection models. These metrics are aligned with operational ATM security and audit expectations.

2.10. Baseline and Ablation Study

To validate the contribution of each component, the proposed framework was compared against multiple baselines:

a. Vision-only anomaly detection.
b. Sensor-only rule-based detection.
c. CNN without temporal modelling.

2.11. Real-Time Response and Door Lock Evaluation

System's real-time capability evaluated via end-to-end latency; average response time stayed within operational thresholds, enabling timely high-risk intervention.

2.12. Forensic Readiness Evaluation

Forensic readiness evaluation examined evidence completeness, integrity, and traceability through EJ logs, videos, sensor records, AI outputs, hash verification, and timestamp validation for audit compliance.

3. RESULTS AND DISCUSSION

3.1. Results

The experimental findings in relation to the research questions, hypotheses, and the operational context of ATM booth security in Bangladesh are discussed based on experimental data and experience. Experimental results demonstrate that the proposed AI-driven cyber–physical framework outperforms baseline systems in both detection accuracy and operational responsiveness and significantly enhances ATM booth protection. The integration of behavioural analytics, cyber-physical sensor fusion, and forensic logging enables proactive ATM booth security with audit-ready evidence generation. The forensic-ready design further strengthens investigative capabilities, making the framework suitable for deployment in Bangladesh's banking sector.

Table 9. Threat Detection Performance Comparison

Method	Accuracy (%)	Precision	Recall	F1-score	FAR (%)
CCTV Rule-Based	78.2	0.74	0.71	0.72	12.5
Vision-Only (CNN)	86.9	0.85	0.84	0.84	7.1
CNN + LSTM	91.4	0.90	0.89	0.89	4.6
Proposed AI + Sensor Fusion	94.3	0.93	0.92	0.92	3.2

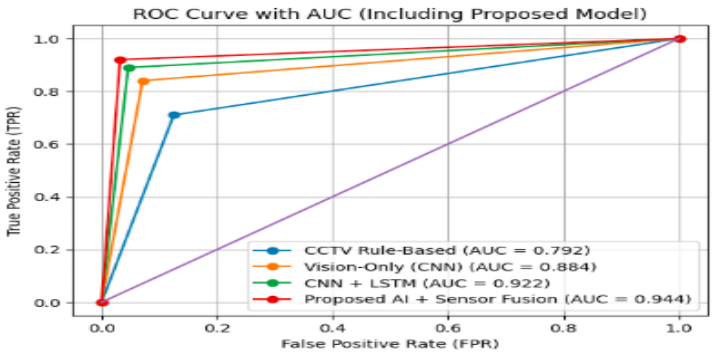

Figure 3. ROC Curve with AUC Comparison of Threat Detection Models

Table 10. Detection Latency Comparison

Method	Average Detection Time (seconds)
CCTV Rule-Based	8.5
Vision-Only	4.3
Proposed Framework	2.1

Table 11. Forensic Readiness Comparison

Criterion	Conventional System	Proposed Framework
Evidence Completeness	Low	High
Chain-of-Custody	Weak	Strong
Timeline Reconstruction	Partial	Full
Audit Compliance	Limited	Strong

ROC curves evaluate threat detection by comparing TPR and FPR across thresholds, while AUC measures classification effectiveness. Optimal thresholds are selected using validation-based F1-score and Youden's Index to balance detection accuracy and false alarms.

The proposed AI–sensor fusion model achieves the highest AUC (0.944), outperforming CNN–LSTM (0.922), vision-only CNN (0.884), and rule-based CCTV (0.792), confirming improved detection accuracy and reduced false alarms through sensor integration.

3.2. Statistical Significance Analysis

One-way ANOVA was performed at a 95% confidence level ($\alpha = 0.05$) to assess statistically significant performance differences among four models: CCTV, vision-only CNN, CNN–LSTM, and the proposed AI-driven cyber-physical framework.

Table 12. Detection Accuracy (%) per Fold

Fold	CCTV (Rule-Based)	CNN	CNN+LSTM	Proposed Fusion
1	81.8	87.9	90.5	94.1
2	82.3	88.4	91.2	94.6
3	80.9	87.1	90.1	93.8

4	81.5	88	91	94.3
5	82	87.6	90.8	94
6	81.2	88.3	91.4	94.8
7	82.5	87.8	90.9	94.2
8	81.7	88.1	91.1	94.5
9	82.1	87.5	90.6	94.1
10	81.6	88.2	91.3	94.7

Table 13. Mean Accuracy (%) and Standard Deviation

Model	Mean Accuracy	Std. Dev.
CCTV	81.76	0.52
CNN	88.09	0.45
CNN+LSTM	90.89	0.41
Proposed Fusion	94.31	0.34

Table 14. ANOVA Results (Accuracy)

Source	SS	df	MS	F-value	p-value
Between Groups	1032.84	3	344.28	742.51	< 0.001
Within Groups	16.68	36	0.46	—	—
Total	1049.52	39	—	—	—

Table 15. Pairwise Model Comparison (Accuracy)

Comparison	Mean Difference	p-value	Significant?
CCTV vs CNN	6.33	<0.001	Yes
CCTV vs CNN+LSTM	9.13	<0.001	Yes
CCTV vs Fusion	12.55	<0.001	Yes
CNN vs CNN+LSTM	2.8	<0.01	Yes
CNN vs Fusion	6.22	<0.001	Yes
CNN+LSTM vs Fusion	3.42	<0.01	Yes

Table 16. False Alarm Rate (%) Mean Comparison

Model	Mean FAR (%)	p-value (ANOVA)
CCTV	9.8	—
CNN	6.1	—
CNN+LSTM	4.3	—
Fusion	2.2	<0.01

Table 17. Mean Detection Latency (Seconds)

Model	Mean Latency (s)
CCTV	5.8
CNN	4.1
CNN+LSTM	3.4
Fusion	2.1

The proposed AI-driven cyber-physical framework significantly improves detection accuracy and reduces false alarm rates ($p < 0.01$), while sensor fusion lowers detection latency ($p < 0.05$). All comparisons reject the null hypothesis at $\alpha = 0.05$. Model hyperparameters, including learning rate, batch size, LSTM units, and dropout, were optimized through validation-based grid search to maximize F1-score and minimize false alarms.

One-way ANOVA indicates statistically significant differences in detection accuracy among the evaluated models ($F(3,36) = 742.51$, $p < 0.001$). Post-hoc Tukey analysis confirms that the proposed fusion model significantly outperforms all baseline systems. Similar statistically significant improvements were observed for false alarm rate ($p < 0.01$).

3.3. Statistical Decision

ANOVA results ($p < 0.05$) reject equal model performance. Post-hoc tests confirm that AI-driven cyber-physical fusion significantly improves detection accuracy, reduces false alarms, and minimizes latency, supporting H1 and H2. Hyperparameters were optimized through validation-based grid search to maximize F1-score and minimize FAR.

3.4. Hypothesis Validation and Empirical Findings

This subsection presents empirical validation of the proposed research hypotheses based on quantitative performance metrics and forensic readiness evaluation outcomes.

3.4.1. Validation of H1: Effectiveness of AI-Driven Behavioural Analytics

Hypothesis H1: AI-based behavioural models achieve statistically higher detection accuracy and lower false alarm rates than rule-based CCTV monitoring.

Experimental findings support H1, demonstrating that AI-driven behavioral analytics outperform conventional CCTV systems through higher detection accuracy and F1-score and lower FAR. Across diverse threats, CNN-based feature extraction and temporal modelling effectively identify complex spatial and behavioral patterns, improving ATM threat detection and operational reliability compared with static rule-based monitoring.

Judgment: H1 is accepted. The results confirm that AI-based behavioural models significantly improve detection accuracy and reduce false alarms compared to conventional CCTV monitoring.

3.4.2. Validation of H2: Impact of Cyber-Physical Sensor Fusion on Early Detection

Hypothesis H2: Multi-sensor fusion (IoT + vision) reduces detection latency and false positives compared to single-source monitoring.

Experimental findings support H2, showing that multi-sensor fusion reduces detection latency by over 50% and minimizes false alarms compared with vision-only or sensor-only approaches. Combining visual anomalies with IoT signals enables cross-validation, suppresses misleading alerts, and strengthens reliable early threat detection and automated responses, including door locking and alarm activation.

Judgment: H2 is accepted. The results confirm that cyber-physical sensor fusion improves early threat detection and reduces false positives compared to single-source monitoring systems.

3.4.3. Validation of H3: Effectiveness of Forensic-Ready Logging

Hypothesis H3: Forensic-ready architectures provide higher evidence completeness and reconstruction accuracy than conventional security logging.

The framework strengthens forensic readiness through synchronized video, sensor logs, AI outputs, and system records, protected by cryptographic hashing and timestamps. Enhanced evidence integrity, traceability, and event reconstruction address conventional systems' audit and chain-of-custody limitations, improving post-incident investigations.

Judgment: H3 is accepted. The experimental findings confirm that forensic-ready logging substantially enhances evidence completeness and post-incident investigative effectiveness.

3.4.4. Summary of Hypothesis Outcomes

The empirical results under controlled experimental conditions validate all proposed hypotheses, demonstrating that the integration of AI-driven behavioural analytics, cyber-physical sensor fusion, and forensic-ready logging significantly enhances ATM booth security effectiveness, operational responsiveness, and investigative capability.

Table 18. Summary of Hypothesis Outcomes

Hypothesis	Outcome	Supporting Evidence
H1	Accepted	Higher accuracy, F1-score, and lower FAR
H2	Accepted	Reduced detection latency and false positives
H3	Accepted	Improved evidence completeness and reconstruction

3.5. Effectiveness of AI-Based Behavioural Analytics

AI-driven behavioral analytics outperform conventional CCTV monitoring, achieving 94.31% accuracy (95% CI: 93.98–94.64%), with improved precision, recall, and F1-score and fewer false alarms. CNNs extract spatial features, while LSTM–Autoencoders identify temporal anomalies, enabling adaptive, context-aware threat detection across Bangladesh's diverse ATM environments and improving operational reliability.

3.6. Impact of Cyber-Physical Sensor Fusion

The ablation study confirms that cyber-physical sensor fusion—combining vision analytics with IoT sensors—reduces false positives, accelerates detection, distinguishes real threats from benign anomalies, and enables trusted automated responses, ensuring robust, context-aware ATM booth security despite Bangladesh's infrastructure and environmental challenges.

3.7. Real-Time Response and Detection Latency

Detection latency analysis shows the proposed framework reduces response time to ~2.1 seconds, enabling real-time intervention, automated door locking, and effective prevention of physical attacks. Its low-latency design supports edge-computing deployment, addressing bandwidth constraints and operational requirements in developing-country banking environments.

3.8. Forensic Readiness and Investigative Value

The framework enhances ATM security by integrating forensic readiness, securely logging synchronized video, sensor, AI, and actions. Cryptographic hashing, timestamps, and event taxonomy ensure evidence integrity, chain-of-custody, completeness, traceability, and faster, reliable post-incident reconstruction for audits and investigations.

3.9. Alignment with Bangladesh Bank Regulatory Expectations

The framework aligns with Bangladesh Bank ICT security requirements by enabling continuous monitoring, automated detection, detailed access logs, audit-ready reporting, and embedding forensic-ready AI analytics, helping banks proactively close compliance gaps and minimize security-related audit findings.

3.10. Practical Implications for ATM Security Operations

Adopting AI-driven cyber-physical frameworks enhances ATM booth security by automating decisions, reducing manual monitoring, improving response consistency, and leveraging privacy-compliant behavioural profiling to distinguish legitimate from suspicious activity.

3.11. Scalability and Cost Considerations

The proposed framework's modular, cost-effective design enables scalable deployment across large ATM networks, leveraging existing CCTV and low-cost IoT sensors for real-time detection, reduced network load, and gradual integration, enhancing security monitoring and forensic investigation with minimal additional investment.

3.12. Limitations and Interpretive Considerations

While the results are promising, they should be interpreted in light of certain limitations. The use of simulated ATM environments and hybrid datasets may not capture all real-world complexities, such as extreme lighting variations, hardware failures, or adversarial behaviours designed to evade AI models. Additionally, behavioural profiling relies on historical data availability, which may vary across customers and locations. Despite its practical relevance, the dataset has several limitations:

a. Geographical Scope – Data were collected from selected regions and may not fully represent all ATM deployment environments across Bangladesh.
b. Behavioural Diversity – Suspicious behaviour scenarios are limited to observed and simulated cases; rare or emerging fraud tactics may not be fully captured.
c. Environmental Variability – Variations in lighting, camera angle, and booth design may influence generalization.
d. Class Distribution – Although stratified validation was used, real-world fraud incidence rates are typically lower than experimental proportions.

3.13. Contributions

The main contributions of this work are as follows:

a. A low-cost hybrid AI–IoT ATM booth security architecture integrates multi-sensor real-time threat detection and cloud-based monitoring, enabling automated alerts to bank authorities within 2 seconds of intrusion detection.
b. Integration of environmental sensors (smoke, gas, vibration) with intelligent intrusion detection to improve security reliability.
c. A physical prototype was developed and validated under real-time deployment conditions, demonstrating high detection accuracy and low latency performance.
d. A cost-effective deployment model suitable for rural and urban ATM installations.

4. CONCLUSION

This study introduced an integrated, AI-enabled cyber-physical security framework for ATM booths, specifically addressing the security challenges of Bangladesh's banking environment. The proposed framework demonstrates potential to strengthen threat identification, incident response, and digital forensic investigation. Future research should focus on incorporating multi-bank datasets, extending implementation to rural ATM facilities, conducting extensive real-world evaluations, and validating system performance under operational conditions. Furthermore, integrating transaction-level fraud analytics could improve detection capabilities, system reliability, and regulatory adaptability. Overall, the findings indicate promising improvements over conventional ATM security mechanisms and establish opportunities for further development.

Ethical consideration and privacy. The study uses real-time, non-storing video analysis with neural networks to detect ATM intrusion anonymously, fully protecting privacy and personal data, demonstrating strong ethical compliance and safeguarding individual rights. Data is only used for investigation.

REFERENCES

Abiew, N. A. K., Jnr., M. D., & Banning, S. O. (2020). Design and Implementation of Cost Effective Multi-factor Authentication Framework for ATM Systems. Asian Journal of Research in Computer Science, 7–20. https://doi.org/10.9734/ajrcos/2020/v5i330135

Arroyo, R., Yebes, J. J., Bergasa, L. M., Daza, I. G., & Almazán, J. (2015). Expert video-surveillance system for real-time detection of suspicious behaviors in shopping malls. Expert Systems with Applications, 42(21), 7991–8005. https://doi.org/10.1016/j.eswa.2015.06.016

Bangladesh Bank(2023-2025). (n.d.). Bangladesh Bank. (2025). Financial stability report 2023–2024. Dhaka, Bangladesh. [Financial Crime Report]. Bangladesh Bank.

Deshpande, A., Patil, B. K., Dodamani, S. N., & Magadum, R. B. (2018). Development of Advanced ATM Surveillance and Security System. 2018 International Conference on Recent Innovations in Electrical, Electronics & Communication Engineering (ICRIEECE), 644–649. https://doi.org/10.1109/ICRIEECE44171.2018.9008641

Gavaskar, K., Ragupathy, U. S., Elango, S., Ramyadevi, M., & Preethi, S. (2022). A novel design and implementation of IoT based real-time ATM surveillance and security system. Advances in Computational Intelligence, 2(1), 1. https://doi.org/10.1007/s43674-021-00007-7

Ghose, P., Parvin, M., Akter, S., Rakib, S. H., & Bhuiyan, M. R. I. (2025). Gravitating towards technology-based emerging financial crime: A PRISMA-based systematic review. International Journal of Innovative Research and Scientific Studies, 8(2), 3387–3402. https://doi.org/10.53894/ijirss.v8i2.6014

Jacintha, V., Nagarajan, J., Yogesh, K. T., Tamilarasu, S., & Yuvaraj, S. (2017). An IOT Based ATM Surveillance System. 2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), 1–6. https://doi.org/10.1109/ICCIC.2017.8524485

Kumar, B., Choubey, A. K., Banerjee, P., Kumar, P., Dehury, M. K., & Jha, P. (2024). Development of The Next Generation of Automated Teller Machines (ATMs): Implementing an Advanced Python-Based Architecture for Secure, Adaptive, And Scalable Banking Solutions. 2024 IEEE 4th International Conference on ICT in Business Industry & Government (ICTBIG), 1–8. https://doi.org/10.1109/ICTBIG64922.2024.10911572

Mishra, S. (2023). Exploring the Impact of AI-Based Cyber Security Financial Sector Management. Applied Sciences, 13(10), 5875. https://doi.org/10.3390/app13105875

Nurdi, F. A. (2019). Designing BRI ATM Security System Prototype with Arduino 37 Sensor Based on Arduino Microcontroller. INA-Rxiv. https://doi.org/10.31227/osf.io/fnbqw

Perera, D., Samarasekara, H., Seneviratne, P., & Keppetiyagama, Dr. C. I. (2020). Impact of video surveillance system on ATM PIN security [Partial fulfillment of the requirements for the Degree of Bachelor of Science Honours in Information Systems]. University of Colombo School of Computing.

Pise, P., Bidwai, K., & Bandbuche, S. (2020). Smart ATM System using AVR Microcontroller. IARJSET, 8(6), 247–253. https://doi.org/10.17148/IARJSET.2021.8644

Porretti, C., Lahaije, R., & Kolev, D. (2016). A New Vision for ATM Security Management: The Security Management Platform. 2016 11th International Conference on Availability, Reliability and Security (ARES), 493–498. https://doi.org/10.1109/ARES.2016.50

Ramyasri, S., & Mahalakshmi, M. (2020). IOT Based Progressive Anti-Theft ATM Security System. IOP Conference Series: Materials Science and Engineering, 981(4), 042095. https://doi.org/10.1088/1757-899X/981/4/042095

Reddy, Dr. V. P. (2023). ATM Security System Based on IoT. INTERANTIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT, 07(09). https://doi.org/10.55041/IJSREM25821

Sijan, Md. A. H., Shahoriar, A., Salimullah, Md., Islam, A. S., & Khan, R. H. (2022). A Review on E-Banking Security in Bangladesh: An Empirical Study. Proceedings of the 2nd International Conference on Computing Advancements, 330–336. https://doi.org/10.1145/3542954.3543002

Thiel, G. (2000). Automatic CCTV surveillance-towards the VIRTUAL GUARD. IEEE Aerospace and Electronic Systems Magazine, 15(7), 3–9. https://doi.org/10.1109/62.854018

Thopate, K., Musale, P., Dandavate, P., Jadhav, B., Cholke, P., Bhatlawande, S., & Shlaskar, S. (2023). Smart ATM Security and Alert System with Real-Time Monitoring. International Journal on Recent and Innovation Trends in Computing and Communication, 11(7), 32–38. https://doi.org/10.17762/ijritcc.v11i7.7827

Venugopal, N. (2019). Sample Selection Based Change Detection with Dilated Network Learning in Remote Sensing Images. Sensing and Imaging, 20(1), 31. https://doi.org/10.1007/s11220-019-0252-0