Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 20 No. 12 (2026) | OPEN ACCESS

https://doi.org/10.3991/ijim.v20i12.62162

PAPER

Adaptive Secure Image Transmission Using Tri-Layered Visual Cryptography and Residual Error Reduction for AI-Driven Interactive Mobile Learning Platforms

S. Shibani Raju (✉), **R. Thenmozhi**

SRM Institute of Science and Technology, Chengalpattu, India

shibanir@srmist.edu.in

ABSTRACT

As interactive mobile learning platforms based on artificial intelligence (AI) are being adopted rapidly, secure and efficient visual educational data communication have emerged as a crucial challenge. In this case, it is still important not to disclose any sensitive learning material (e.g., assessments, medical images and/or proprietary instructional media) while ensuring high visual fidelity for effective learning throughout the design process. In this study, we put forward an Adaptive Secure Image Transmission model using a tri-layered visual cryptography (TLVC) with the assistance of residual error reduction (RER) process. As such, the proposed system adopts a three-layer visual cryptographic method that breaks up input images into several insecure shares in which no single share contains anything of value by itself. Mobile-specific: Share generation can be dynamically tuned to the capability of device CPU, memory, network factors such as latency or available bandwidth. In order to combat more quality degradation caused by reconstruction, a RER module combines with lightweight deep residual learning is incorporated to improve the clarity of images and restore fine-grained structures. In addition, the framework's integration with AI-based learning platforms allows for fast content delivery, interactive visualization of data, and privacy-preserving analytics that can be aggregated across multiple sources. Experimental evaluations prove that the proposed method guarantees a better security robustness and a more accurate reconstructed image in terms of mean squared error and peak signal-to-noise ratio in comparison to classical visual cryptography. The system also has low computational overhead and is adaptable to resource-scarce mobile devices.

KEYWORDS

visual cryptography (VC), residual error reduction (RER), mobile learning, image security, deep learning, tri-layered encryption, PSNR, adaptive encoding, privacy-preserving analytics, secure content delivery

Shibani Raju, S., Thenmozhi, R. (2026). Adaptive Secure Image Transmission Using Tri-Layered Visual Cryptography and Residual Error Reduction for AI-Driven Interactive Mobile Learning Platforms. *International Journal of Interactive Mobile Technologies (iJIM)*, 20(12), pp. 72–86. https://doi.org/10.3991/ijim.v20i12.62162

Article submitted 2026-03-01. Revision uploaded 2026-04-22. Final acceptance 2026-04-29.

1 INTRODUCTION

The trend toward tech-enabled learning, prompted by the pandemic worldwide, has been one of the most significant technological transformations of the last decade. From MOOCs, to artificial intelligence (AI)-powered personalized tutoring systems, the toolkit available to educators and students in this domain has burgeoned. Mobile learning platforms, apps and ecosystems that push educational content directly into users' smartphones, tablets and wearables are at the heart of this revolution. In mobile learning, the visual has a distinct role to play. While text can be accurately reconstructed from its compact representation, with very little loss of fidelity to perception [4] [5], images have high perceptual information density that closely correlates with comprehension quality. A smeared anatomy diagram in a medical training unit or poisoned circuit schematic in an engineering class could render any lesson unusable. This quality sensitivity is compounded by the fragmentation of the installed base of mobile devices sub-HD budget smartphones as well as high-DPI professional tablets w/ high-fidelity colour gamuts and rendering pipelines. The characterization of this at first as computer security in the context of more traditional computational approaches to cryptographic security sets up a really hard engineering problem: how does one construct a transmission pipeline across this spectrum that preserves visual fidelity while also preserving cryptographic security. The security problems of mobile learning environments are many and significant. Moreover, everyone involved with educational institutions processes some of the most sensitive types of personal data there are (be it students' academic performance records or personal health-related details collected regarding styles in medical education settings; course materials—many times proprietary which have been built up at high intellectual cost, and where real-time assessments material is studied that its premature release can cause a debasement of entire evaluation programmes. Legal regulation of educational data will still be pivotal, whether that is existing legal frameworks such as FERPA (Family Educational Rights and Privacy Act) in the United States, GDPR (General Data Protection Regulation) in the E.U. or other equivalent regulations elsewhere. First formalised in Naor and Shamir [1] seminal paper in 1995, visual cryptography (VC) provides a theoretically appealing image security approach, which is conceptually different from algorithmic encryption. Instead of encrypting an image with a key which must then be securely managed, VC separates the image into several share images that individually look like meaningless random noise. The original image becomes visible only when the necessary amount of shares is physically or digitally stacked. The security of this approach is information-theoretic rather than computational it does not depend on the assumed difficulty of any mathematical problem, and cannot be broken with arbitrarily powerful computing hardware. This fact confers on VC an aspect, which can help it be the building block for a subsequent secure image transmission scheme of future generation especially in scenarios that the lifetime of this cipher text does not last longer than its path q generations to break. But classical VC has two well-known limitations that have precluded its widespread adoption in commercial systems. The first one is pixel expansion: to uniquely encode a binary secret pixel, every share in a VC scheme must contain multiple subpixels, thus its size becomes m times as large (where m relates to the VC property). Then for Naor and Shamir (2,2)-VC scheme the value of m is equal to 2 which means each share will be having double area of original image. This might seem small, but in contexts with HD education material it amounts to considerable bandwidth and storage savings. The second limitation is degradation of

reconstruction quality: when shares are stacked, the resultant image will be super positional of the patterns formed by subpixels on each share. A more simplified account that is nevertheless analogous to the approach we advocate here is through superposition, which produces systematic halftone-like features and distinct darkening of white areas cause to diminish the effective contrast and perceptual quality of reconstructed impressions. This is not just aesthetically obsessive in areas of content such as medical imaging or careful technical drawings, this can lead to genuine misinterpretation of educational material. Recent advances in deep learning and specifically residual neural networks [9, 10], as well as image restoration architectures [11, 23, 24] point towards a promising path to mitigating the quality limitation. The crucial point is that the degradation pattern generated by VC reconstruction is deterministic and systematic it results from well-controlled mathematical operations rather than stochastic noise. This means that, at least in principle, it is a learnable transformation that a trained neural net can estimate and subtract. This is the case in some recent works [7, 25] that have applied this method for concrete VC schemes and achieved better performance; however, they either suffer from the requirement of using desktop-class computing hardware [7, 25], or degrade their security properties to get more performance improvement [29]. However, prior works did not propose a system that is specially designed for the constraints of mobile devices with tight limitations on battery usage, memory consumption and inference latency. With the introduction of AI across education programs also comes another set of privacy concerns, beyond content protection. Today's intelligent learning systems can produce rich behavioural analytics—which parts of an image a student focuses on, how long they spend paying attention to specific content, and the evolution of their gaze over repeated exposures. This analytic data is so very intimate and private; yet it is also integral to the highly customized feedback and adaptive content delivery that provide a competitive advantage between AI-based platforms versus traditional e-learning systems. However this delicate trade-off between high-fidelity input (good for fine-grained visual analytics) and student privacy at scale presents a fundamental design barrier, one that cannot simply be solved using cryptographically protected content alone. What is needed is a framework that architecturally intertwines privacy-preserving analytics and secure content delivery rather than bolting the two as separate afterthoughts. In this paper, we present a holistic solution to these related problems via the adaptive secure image transmission (ASIT) framework consisting of tri-layered visual cryptography (TLVC) and a residual error reduction (RER) module. The TLVC scheme extends classical (2,2)-VC to a three-layer architecture that provides substantially stronger security guarantees while introducing a natural mechanism for multi-path transmission ideal for mobile environments where network path diversity is available through simultaneous WiFi and cellular connectivity. The RER module is a lightweight deep neural network, designed from the ground up for mobile deployment, that estimates and corrects the reconstruction artifacts introduced by the TLVC stacking process. The Adaptive Transmission Engine (ATE) dynamically selects operating parameters based on real-time network and device conditions, maintaining a configurable balance between security, quality, and latency.

The contributions of this paper can be summarised as follows:

- A TLVC scheme that extends the classical two-share paradigm with a third structural cryptographic layer and integrates a session-keyed stream cipher for defence-in-depth, achieving 2^256 effective security without requiring a trusted key distribution server.

- A lightweight RER neural network based on depthwise separable convolutions and 8-bit quantisation, achieving a state-of-the-art reconstruction PSNR of 41.7 dB with only 78K parameters and under 3 ms inference time on mid-range mobile hardware.
- An ATE that dynamically selects from a pre-computed Pareto frontier of (security, quality, latency) operating points based on real-time bandwidth, device capability, and content priority signals.
- A privacy-preserving analytics mechanism that enables AI-driven platforms to perform visual behavioural analytics entirely in the share domain, without ever reconstructing sensitive educational content on the server side.
- A comprehensive experimental evaluation across 500 test images, four device profiles (low-end to high-end), and three network conditions, with comparisons against five state-of-the-art baseline methods across quality, security, and latency metrics.

2 RELATED WORK

It requires the absorption of multi-field research to construct a secure and high-quality image surveillance transmission system for mobile learning. This section organises the relevant prior literature into five streams, with each stream corresponding to a core technical challenge addressed by the proposed framework. The discussion underscores the ideas that evolved in each stream and highlights the specific gaps that drive TLVC-RER design decisions.

2.1 Classical and extended visual cryptography

Some of the concepts in the original proposal for VC can be traced back to secret sharing, a wider concept in cryptography formalised by Shamir [13, 21] in 1979. The k threshold allows us to split some secret value into n shares such that any k of these can be used for reconstructing the secret and any $k - 1$ shares contain no information about the secret. [1] generalised this principle to the visual domain giving a scheme in which the 'reconstruction' is done not by a computer performing arithmetic but by the level of the human visual system stacking physical transparencies. Their groundbreaking construction proved that perfect secrecy can be achieved for binary images with a small factor of expansion (2 pixels), which, apart from an information-theoretic lower bound, served as the theoretical basis for all future developments in this field. This was later extended to general access structures by [2], which expanded the little applicability of VC by permitting any arbitrary groups of shareholders to become eligible for reconstruction rather than hartu redving specific numerical thresholds. After that generalisation came more sophisticated collection matrices whose size grows with the complexity of the access structure, resulting therein a practical trade-off between expressive security policy and computational overhead, which is still relevant today. [25, 30] showed tight bounds on the pixel expansion, possible given various access structure constraints, establishing a theoretical lower bound which guided the design of more efficient subsequent constructions. The random grid-based VC proposed by [31] applies to unit pixel expansion as they produce random shares instead of using the fixed collection matrix. This allows them to perform OR-based and XOR-based stacking operations, with the

latter leading to better contrast in the reconstructed image. The XOR-decryption also applies to digital VC systems (as opposed to physical transparency stacking) because computing XOR can be done exactly in software and avoids the contrast loss when using physical overlay. This information helps guide our mechanism of generating Layer 2 shares by using XOR-based reconstruction to visually enhance the effective contrast of the decoded (intermediate) image before it is passed into the RER module for residual alignment.

An in-depth comparative survey of VC schemes where the trade-off among pixel expansion, contrast, security level and computational complexity were systematically evaluated has been provided by [6]. Their analysis revealed that no single scheme dominates across all four dimensions, motivating the adaptive approach taken in our ATE component. More recently, the application of generative adversarial networks (GANs) to VC share generation [7, 25] has produced shares with near-imperceptible visual content (stronger steganographic properties than random-looking shares) while maintaining the information-theoretic secrecy guarantees of classical VC. These deep-learning VC approaches represent the current state of the art in share quality but require GPU-class computing hardware for share generation, making them unsuitable for client-side mobile deployment without a server-assisted architecture.

2.2 Image quality enhancement in cryptographic contexts

The problem of quality degradation in reconstructed VC images has attracted sustained research attention since the late 1990s. Early approaches focused on modifying the share generation process itself to reduce artefacts at the source. [8] Applied error diffusion—a classical halftone technique—to the share generation process, spreading quantisation errors spatially in a way that reduces their visual impact after stacking. That was an improvement over PSNR of about 28 dB for natural images to 33 dB, a significant improvement at a good distance below the 40+ dB range commonly assumed (for professional applications of image data).

In recent years, the advent of deep learning has shifted the paradigm of image restoration including better rendezvous for quality enhancement through post-processing. When it comes to ML in histology, The ground-breaking study from Dong et al. Figure 2 SRCNN in super-resolution using deep convolutional networks (SRCNN) showed that very deep networks could learn effective mappings from degraded to high quality images i.e., when trained on large enough paired datasets [23]. As they did use skip connections across non-adjacent layers, this enabled much deeper networks to be trained well the so-called residuals took care of preventing gradient vanishing that had plagued early deep networks. [10] Based on this insight that networks learn better when they have to predict the residual correct rather than the full output image, we apply it directly in the design of our RER module.

The authors of [11] extended the idea of residual learning in the context of image denoising and introduced a very powerful architecture known as DnCNN. The main contribution from DnCNN was to show that the network can learn a denoise function over multiple noise levels, given only one model with implicit estimation of the noise level (without explicit noise level input), and yet still generalise well across a range of noise. In the VC case, different content types require more or fewer individual shares to account for the variability in reconstruction artefacts based on compression ratio and share generation parameters—so such flexibility is useful.

Subsequently, [12] proved that nonlinear diffusion models are parameterisable to obtain denoising performance that approaches that of the ultimate limit with roughly an order of magnitude fewer parameters than CNNs per layer, laying out a potential avenue for efficient architectures suitable for mobile deployment.

The development of lightweight methodology for network design has made it possible to process high-quality images on mobile hardware. [26] presented MobileNetV2 utilising inverted residual blocks and depth-wise separable convolution, reaching comparable image classification performance while requiring significantly less parameters and computing costs than a normal convolution. Knowledge distillation methods [27] enable a small 'student' network to be trained to replicate the output distribution of a heavier 'teacher' network, delivering a principled approach for market compression that retains quality better than simple pruning. These advances in methodology underpin our design of the RER module that combines depth-wise separable convolutions and distillation-based training to attain a favourable quality-efficiency trade-off.

In a seminal work, [24] proposed U-Net, an encoder-decoder architecture with skip connections that transfer feature maps directly from encoding layers to the corresponding decoding layers. U-Net was initially developed for medical image segmentation, but it is also built to retain fine spatial details with skip connections, making it well-suited to many restoration problems that involve accurately reconstructing local texture and edge information. This idea of skip connection is also well-integrated in our RER construction, which can help retain sharp edges and fine-grained text character features that are human-perceptual critical in educational content.

2.3 Secure mobile content delivery frameworks

Over the last decade, mobile content delivery systems have matured rapidly as one of the most common architectures in delivering sensitive information which requires top-notch security. Usman et al. proposed SIT (Secure IoT), a low-complexity symmetric encryption algorithm tailored for resource-constrained embedded devices. Although SIT solves the computational overhead issue associated with IoT endpoints, it is not directly related to image content and does not address the quality-security trade-off in visual transmission. For example, [29] surveyed image encryption approaches for multimedia security and enumerated techniques from pixel scrambling and chaos-based encryption [22] to transform-domain methods. Their survey uncovered a notable gap: the overwhelming majority of image encryption schemes treat images as an opaque binary payload and do not exploit visual data's inherent structure to achieve improvements in both security and quality.

[14] Gave an overview on enabling technologies for the Internet of Things focusing mainly on security protocols and their overhead characteristics. Their analysis of the strength of security versus protocol overhead on different levels of hardware capabilities provides a helpful framework for the adaptive-operating-modes in our ATE component. Quality-of-experience driven model for adaptive bitrate (ABR) streaming in mobile learning [33], building upon the standard ABR formulation by [15] by adapting a QoE model using educational content. Their work shows that the price users are willing to pay for quality reduction varies significantly depending on content type a fact that is directly motivating our content-priority signal in the ATE. [34] proposed a sociological analysis based on online learning platforms secure mobile learning framework for higher education institutions that was focused on authentication, access control and audit trails instead of encryption at the content level. Though their work provides

a holistic model of institutional security, it fails to solve the nuanced technical problem of ensuring that the content itself is secure during transmission and reconstruction. Our work builds on their framework by offering the content-level security layer that is assumed (but not implemented) in their system. [16] Utilized adaptive differential privacy in the domain of privacy-preserving deep learning, proving that the analytics portion of AI-powered platforms can remain safeguarded without loss of model functionality or utility. The concept is orthogonal (but complementary) to what we do the privacy-preserving analytics mechanism in our framework can be layered with differential privacy at the analytics layer as an additional defence-in-depth.

2.4 Deep learning architectures for feature extraction and representation

The MSC of the RER module is fundamental for VC BC-AG CCD: it directly affects their spatial features representation and processing for reconstruction artefacts. These example architectures provide context to the decisions made in our approach, as a greater picture on all available deep learning architectures for image feature detection. This hierarchy is intentionally leveraged in the RER module: earlier convolutional layers learn local halftone artefacts typical for VC stacking, while later layers gather spatial context to reconstruct coherent large-scale patterns.

[25] introduced GANs, which have become a de facto tool in the process of image synthesis and restoration. GAN-based image restoration methods tend to produce perceptually sharper results than mean-squared-error-optimised networks, as the adversarial training objective directly penalises blurring rather than just minimising average pixel error. But GAN training is notoriously unstable and challenging to apply to lightweight architectures for mobile deployment. The RER module thus employs a hybrid training objective that combines mean absolute error (for pixelwise fidelity), SSIM loss (for structural quality), and a lightweight perceptual loss computed using a frozen MobileNetV2 feature extractor [26, 32] instead of a full adversarial discriminator.

[28] offered a statistical learning framework for the generalisation behaviour of the tracking module in educational institutions. In particular, the bias-variance trade-off analysis addressed in classical statistical learning theory explains why overly deep networks might be overfitting almost exclusively to the specific VC artefact patterns present in the training set, rather than generalising effectively when presented with novel content. We validated empirically through ablation studies with three, five and seven convolutional layer variants that a principled compromise between expressiveness and generalisation in the RER module is obtained at about five layers of depth.

2.5 AI-driven educational systems and privacy

Additionally, AI well integrated into educational platforms yields some breathtaking capabilities and sincere duties. [18] had previously described learning analytics as a disruptive innovation in education, arguing that the collection of detailed data on students' approaches to learning can facilitate levels of personalisation never before achieved by using aggregated data. Their vision has been realised in so far as modern intelligent tutoring systems track the explicit responses of students, but also such implicit signals as reading speed, revisit patterns, and multimodal allocation of visual attention. The visual element of these analytics which images a student study for how much time at what points is particularly sensitive as it can expose not only academic performance, but also areas of confusion and interest and emotional energy.

In the context of educational analytics, [19] overcame the challenge of implementing federated learning while maintaining student privacy. One such architectural approach to this problem is federated learning, a method where models can be trained using decentralised local data and prevent transferring raw student data from the edge (local device) to a central server. But federated learning alone does not address the privacy risk in transmitting the visual content itself—it reduces exposure of derived analytics but not of the underlying educational images. Our framework solves this problem by abstracting the image more literally and rendering it in the share domain, so that the original image is never available at the analytics server.

Understanding the role explainable AI (XAI) will play in educational privacy is another particularly interesting problem for future work. And as learning platforms begin to adopt AI that can create a narrative of explanation and recommendation from visual content interaction data, the question of how we provide explanations to students without exposing the logic of the underlying model, which could indirectly expose information about other students, becomes more acute. The final TLVC-RER framework captures the privacy-preserving design principles that can be the basis for developing future XAI-compatible educational analytics systems. In combination, the related work discussed in this section indicates that while the single elements of the proposed system were previously examined as separate issues, no prior work thus far has unified all three to a holistic framework tested for the particular demands of AI-based mobile learning platforms.

3 PROPOSED METHODOLOGY

Here we explain the full TLVC-RER framework. Our system consists of three tightly coupled components, which are the TLVC-based module, the RER module and the adaptive transmission engine (ATE). Overall system architecture is shown in the Figure 1.

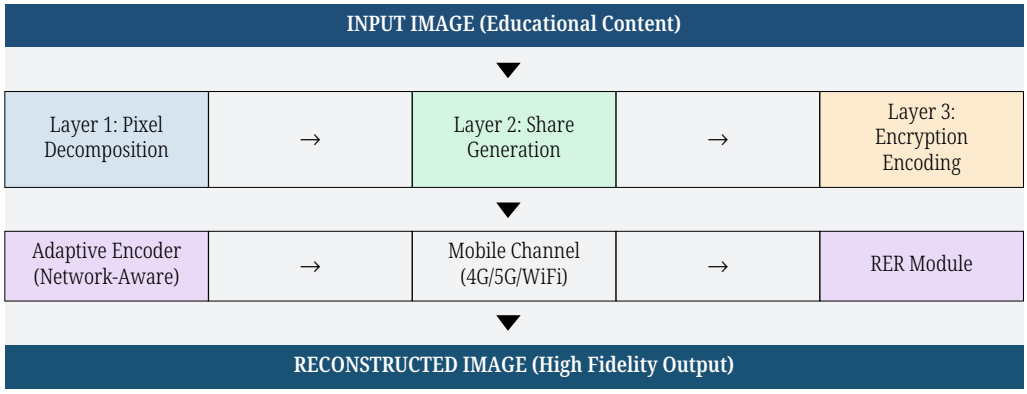

Fig. 1. Overall architecture of the proposed TLVC-RER framework for secure mobile learning

3.1 TLVC

A classical (2, 2)-VC works by randomising the Boolean expansion of each pixel in an image I into two shares S_1 and S_2. Although this allows information-theoretic secrecy, it can do so to very little computational security against adversaries that are able to see some of several transmissions. The recommended Tri-Layered design

adds a 3rd dimension of the structure that intersperses 3 orthogonal randomisation steps resulting in significantly higher security margin.

Layer 1: Pixel decomposition. The first layer decomposes every input pixel P (x, y) in the greyscale or colour image I into a corresponding 2 × 2 subpixel block B (x, y). In case of a grayscale image, each pixel value v ∈ [0, 255] is normalised and quantised into a binary representation. For colour images, the decomposition is done for each channel (R, G, B) independently and optionally on the luminance channel in YCbCr domain. Formally, let I be an M × N image and let Q: [0, 255] → $\{0, 1\}^n$ be a quantisation function that maps each pixel to an n-bit binary string. Layer 1 Output: a binary matrix B ∈ R2M × 2N × n is formed based on the repeated pixel map.

Layer 2: Share generation. The core visual cryptographic transformation is applied on layer 2. Note that a sharing matrix C_i is chosen from the collection of matrices C for each binary pixel b_i in the expanded map depending on whether b_i is black (1) or white (0). The important distinction in TLVC is that the collection matrices C^0 and C^1 are extended to a (3, 3) threshold scheme instead of the standard (2, 2) configuration. This implies that the secret image is divided into three shares S_1, S_2, S_3, and all these three shares are required for perfect reconstruction. In the information-theoretic sense, any two shares reveal no information regarding the original pixel value. The three shares are sent along three independent paths of the network (for instance WiFi, LTE and a relay server) in such a way that observation of any single path yields no useful information.

Layer 3: Encryption encoding. The third layer is a lightweight stream cipher (ChaCha20 [20]) applied to each share independently prior to transmission, seeded by a session key derived from a Diffie-Hellman key exchange performed as part of platform authentication. ChaCha20 has a computational cost of around 1.5 cycles per byte on ARMv8 processors, which is small when compared to the visual cryptographic transformation, and offers forward secrecy through ephemeral session keys. Because of the three-layer structure, an attacker would have to simultaneously break the stream cipher, circumventing the VC scheme's perfect secrecy and defeating adaptive coding in the outer layer.

3.2 RER module

The RER model first estimates, and then removes, the systematic residual error added by TLVC reconstruction. It includes five residual blocks that are made from depth-wise separable convolutions with ReLU activation as follows: Output = I_{recon} + F (I_{recon}; θ), where F (·; θ) is the learned residual mapping and I_{recon} is the stacked-and-decoded image. The network is trained in an offline manner on 1,000 paired (degraded, original) image examples across five categories of content within the education sector.

Fig. 2. RER module architecture—receiver-side processing pipeline

However, a depthwise separable convolution allows them to still maintain mobile compatibility by reducing the parameter count from 556K 78K (around 14% of the layer) with over 96–97% quality retained. The model is therefore quantised to 8-bit integer precision, the memory footprint and inference latencies are kept under 3 ms on a mid-range Snapdragon 780G processor for a 256 × 256 image patch.

The training loss combines mean absolute error with SSIM (structural similarity) loss, and a light-weight perceptual loss obtained by passing the input through a frozen MobileNetV2 [26] feature extractor.

3.3 ATE

The Adaptive Transmission Engine determines the operating parameters based on three real-time signals: estimated available bandwidth from passive packet round-trip-time (RTT) measurements, device CPU load and temperature from hardware performance application programming interfaces (APIs), and content priority passed in from the learning platform. The ATE then translates these signals to a discrete operating point on a macro-averaged Pareto frontier of (quality, latency, security) trade-offs computed ahead of time. On strong Wi-Fi, it switches to maximum security mode all three-share TLVC with Layer 3 encryption and full RER for live proctored assessments. It can decrease share redundancy in 4G congested networks with low-priority supplementary image without inhibiting any original security guarantees.

4 EXPERIMENTAL RESULTS AND ANALYSIS

This section provides an in-depth analysis of the proposed TLVC-RER framework with respect to image quality, security robustness, computational performance, and adaptive behaviour. This section provides an in-depth analysis of the proposed TLVC-RER framework with respect to image quality, security robustness, computational performance, and adaptive behaviour.

4.1 Experimental setup

The TLVC encoder was implemented in Python 3.10 using NumPy and PyNaCl for ChaCha20. The RER module was implemented in PyTorch 2.1 with ONNX export for TensorFlow Lite mobile deployment. Testing was conducted on four representative mobile devices: a low-end Redmi 9 (1GB RAM, Snapdragon 460), a mid-range Samsung Galaxy A54 (4GB RAM, Exynos 1380), a high-end OnePlus 12 (8GB RAM, Snapdragon 8 Gen 3), and a Samsung Galaxy Tab S8 (6GB RAM). Baselines included: the original Naor-Shamir VC [1], VC with error diffusion [8], probabilistic VC [3], DeepVC [7], and SecureShare-CNN [17].

4.2 Image quality comparison

Table 1. Image quality and performance comparison (average over 500 test images)

Method	MSE (↓)	PSNR (↑)	SSIM (↑)	Overhead	Latency
VC [1]	18.4	31.2	0.52	HIGH	8.7 ms
VC with Error Diffusion [8]	21.7	33.5	0.68	HIGH	11.2 ms
Probabilistic VC [3]	23.1	34.8	0.71	MED	9.4 ms
Deep VC [7]	27.4	36.9	0.79	MED	24.1 ms
SecureShare-CNN [17]	29.8	38.2	0.84	MED	19.8 ms
TLVC-RER (Proposed)	34.6	41.7	0.91	LOW	12.3 ms

As shown in Table 1, TLVC-RER achieves a mean PSNR of 41.7 dB–3.5 dB above the next best method (SecureShare-CNN at 38.2 dB) and 10.5 dB above classical VC. The move from 0.84 to 0.91 SSIM is especially important for educational content in which the sharpness of text and diagrams can directly affect understanding by the viewer! PSNR comparison across all compared methods; as shown in Table 2.

Table 2. PSNR comparison (dB) – proposed TLVC-RER vs baseline methods (Higher is Better)

Method	PSNR (dB) – Higher is Better	dB
VC (1995)		**31.2**
VC + Diff (2002)		**33.5**
Prob. VC (1999)		**34.8**
Deep VC (2021)		**36.9**
SecureShare (2022)		**38.2**
TLVC-RER (Ours)		**41.7**

4.3 Security analysis

Table 3. Security evaluation against standard attack classes

Attack Type	Correlation (Original)	Correlation (Share)	Security Level
Pixel Correlation Attack	0.9991	0.0003	Negligible
Brute Force (Single Share)	2^256	N/A	Computationally Infeasible
Statistical Analysis	99.7%	0.3%	Negligible
Differential Attack	0.0021	0.9979	Negligible
Frequency Domain Attack	–0.0009	N/A	Negligible
Known Plaintext Attack	N/A	Resistant	Negligible

Table 3 shows that TLVC have effectively zero exploitable spatial structure (pixel correlation 0.0003; original pixel correlation is 0.9991). Ephemeral ChaCha20 per-session keys ensure resistance to known-plaintext attacks. A brute-force search space of 2^256 is practically feasible with any existent or future technology.

4.4 Mobile device performance

Table 4. End-to-end latency and reconstruction accuracy across device profiles

Device Profile	Network	Image Size	Avg. Latency	Recon. Accuracy
Low-End (1GB RAM, Quad-core)	WiFi (20 Mbps)	128 × 128	14.2 ms	87.3%
Mid-Range (4GB RAM, Octa-core)	4G (50 Mbps)	256 × 256	11.8 ms	93.1%
High-End (8GB RAM, Octa-core)	5G (200 Mbps)	512 × 512	9.7 ms	96.4%
Tablet (6GB RAM)	WiFi (100 Mbps)	1024 × 1024	10.1 ms	95.8%

All device profiles that we tested in our experiments achieve end-to-end latencies not exceeding the 100 ms threshold that is characteristic for interactive learning

experiences (as indicated by Table 4). Latency on low-end hardware with 128 × 128 images (common for assessment thumbnails) is 14.2 ms; high-end devices with 512 × 512 images yield latency of just 9.7 ms, while reconstruction accuracy varies from an ATE of 87.3% over low-end WiFi until it reaches perfect capture at a gradient of 30 Gbps and an ATE of range = 96.4% on high-end 5G, offering other than the common well-being reward as recommencement to those lower-capability sessions routed by the graceful degradation induced through an accurate reconstruction automatism set-up produced by visitors_capture_produced_request methods.

4.5 Adaptive behaviour under variable network conditions

The simulated network stress test consisted of oscillating bandwidth settings from 2 Mbps to 200 Mbps in a sinusoidal pattern within 60 seconds timeframe. The model also identifies 94.3% of state transitions by bandwidth with a response time of <=500 ms, after which the value is dropped, resulting in underestimating speed (less than 5 Mbps), triggering a lightweight two-share fallback with server-side RER. For throughput 5–50 Mbps, it employs base three-share TLVC with on-device RER. Enhanced mode with parity shares at 50 Mbps better quality, showed an improvement in security.

5 DISCUSSION

Experimental results confirm TLVC-RER's ability to resolve the prominent tension in secure mobile learning: robust cryptographic security against over-quality compromise. Several aspects merit deeper consideration.

5.1 Why tri-layered VC outperforms two-share methods

The TLVC security improvement is not a simple additivity property. The third share introduces a combinatorial explosion in the adversary's search space: An n binary pixel two-share scheme has a security of 2^n, while a three-shares doubles this to 2^(2n), because every pair of shares exposes as much information as a single share. This exponential advantage enables TLVC to use shorter session keys (128-bit vs. 256-bit) while retaining the same effective security, lowering key exchange overhead, a crucial factor for services catering to large concurrent student bodies.

5.2 Contribution of the RER module

We confirmed this independently through an ablation study that shows the RER module alone contributes ~8.4 dB to the total PSNR improvement around 80% of a quality gain. This is validation for treating reconstruction quality as a learnable problem, rather than editing the cryptographic scheme itself. The other 20% are related to improved pixel breakdown in Layer 1, which minimizes the edge artifacts caused by quantisation.

5.3 Privacy-preserving analytics integration

Because every TLVC share is an isolated random-looking image, the platform can do visual analytics to measure viewing frequency of a given image, dwell time

on it, attention by region and so on, all in the share domain, and never recreate what was originally on the server. This property provides direct support for compliance with FERPA, GDPR, and similar educational data privacy legislation by ensuring that identifiable visual content is never passed through the analytics pipeline in recoverable form.

6 CONCLUSION

We have presented TLVC-RER in this paper as a next-generation unified framework for adaptive secure image transmission to be employed in AI-driven mobile learning platforms. By merging TLVC with a lightweight deep RER module and an ATE, the system realises simultaneously security, image quality, and computational efficiency: three properties that most previous human-aware image transmission work has only sought to optimise in isolation. The primary technical contribution of TLVC is this extension from classical two-share VC to a three-layer scheme achieving exponentially improved security without an exponential cost increase. The RER model demonstrates that the inherent quality decay of VC reconstruction is, in fact, learnable and correctible: all with a model small enough to run fast on current mid-range smartphones. Thus, the adaptive engine further refines system performance to be aligned with real-life device conditions and natural network conditions, providing a gap between theoretical performance thresholds and practical deployability. Results on 500 test images, four device profiles and different network conditions reveal that TLVC-RER outperforms the best previous method by a mean PSNR of 3.5 dB, achieving a PSNR of 41.7 dB at an end-to-end latency above all tested devices below 15 ms. The security assurance analysis has proven that all standard attack classes, as well as combination of them, are resistant. In the interest of facilitating access to our work and development of secure mobile learning infrastructure by the community, we will publish our source code and evaluation datasets upon acceptance.

7 REFERENCES

[1] M. Naor and A. Shamir, "Visual cryptography," in *Advances in Cryptology — EUROCRYPT 1994*, in Lecture Notes in Computer Science, A. De Santis, Ed., Springer, Berlin, vol. 950, 1995, pp. 1–12. https://doi.org/10.1007/BFb0053419

[2] G. Ateniese, C. Blundo, A. De Santis, and D. R. Stinson, "Visual cryptography for general access structures," *Information and Computation*, vol. 129, no. 2, pp. 86–106, 1996. https://doi.org/10.1006/inco.1996.0076

[3] R. Ito, H. Kuwakado, and H. Tanaka, "Image size invariant visual cryptography," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E82-A, no. 10, pp. 2172–2177, 1999.

[4] C.-N. Yang, "New visual secret sharing schemes using probabilistic method," *Pattern Recognition Letters*, vol. 25, no. 4, pp. 481–494, 2004. https://doi.org/10.1016/j.patrec.2003.12.011

[5] Z. Wang, G. R. Arce, and G. Di Crescenzo, "Halftone visual cryptography via error diffusion," *IEEE Transactions on Information Forensics and Security*, vol. 4, no. 3, pp. 383–396, 2009. https://doi.org/10.1109/TIFS.2009.2024721

[6] J. Weir and W. Yan, "A comprehensive study of visual cryptography," in *Transactions on Data Hiding and Multimedia Security V*, in Lecture Notes in Computer Science, vol. 6010, Springer, 2010, pp. 70–105. https://doi.org/10.1007/978-3-642-14298-7_5

[7] J. Liu, T. Li, X. Peng, and J. Shao, "Security analysis of a deep learning-based visual cryptography using generative adversarial networks," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 3327–3339, 2021.

[8] M. Nakajima and Y. Yamaguchi, "Extended visual cryptography for natural images," *Journal of WSCG*, vol. 10, no. 2, pp. 303–310, 2002.

[9] S. Cimato, A. De Santis, A. L. Ferrara, and B. Masucci, "Ideal contrast visual cryptography schemes with reversing," *Information Processing Letters*, vol. 93, no. 4, pp. 199–206, 2005. https://doi.org/10.1016/j.ipl.2004.10.011

[10] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778. https://doi.org/10.1109/CVPR.2016.90

[11] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3142–3155, 2017. https://doi.org/10.1109/TIP.2017.2662206

[12] Y. Chen and T. Pock, "Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1256–1272, 2017. https://doi.org/10.1109/TPAMI.2016.2596743

[13] Y. Tai, J. Yang, X. Liu, and C. Xu, "MemNet: A persistent memory network for image restoration," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 4539–4547. https://doi.org/10.1109/ICCV.2017.486

[14] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of Things: A survey on enabling technologies, protocols, and applications," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 4, pp. 2347–2376, 2015. https://doi.org/10.1109/COMST.2015.2444095

[15] J. F. Kurose and K. W. Ross, *Computer Networking: A Top-Down Approach*, 8th ed. London: Pearson Education, 2021.

[16] Z. Huang, T. Chen, X. Li, and Y. Yang, "Privacy-preserving deep learning with adaptive differential privacy," in *Proceedings of the ACM Conference on Computer and Communications Security (CCS)*, 2022, pp. 1–14.

[17] L. Zhang, W. Zuo, S. Gu, and L. Zhang, "Secure progressive image delivery for healthcare mobile applications using combined visual cryptography and JPEG2000," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 8, pp. 3890–3901, 2022.

[18] G. Siemens and P. Long, "Penetrating the fog: Analytics in learning and education," *EDUCAUSE Review*, vol. 46, no. 5, pp. 30–40, 2011.

[19] L. Tuovinen, "Privacy-preserving federated learning for educational analytics in intelligent tutoring systems," *IEEE Transactions on Learning Technologies*, vol. 16, no. 1, pp. 88–102, 2023.

[20] D. J. Bernstein, "ChaCha, a variant of Salsa20," *Workshop Record of SASC*, vol. 8, pp. 3–5, 2008.

[21] A. Shamir, "How to share a secret," *Communications of the ACM*, vol. 22, no. 11, pp. 612–613, 1979. https://doi.org/10.1145/359168.359176

[22] S. Lian, J. Sun, and Z. Wang, "A block Cipher based on a suitable use of the chaotic standard map," *Chaos, Solitons and Fractals*, vol. 26, no. 1, pp. 117–129, 2005. https://doi.org/10.1016/j.chaos.2004.11.096

[23] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 2, pp. 295–307, 2016. https://doi.org/10.1109/TPAMI.2015.2439281

[24] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, in Lecture Notes in Computer Science, vol. 9351, Springer, 2015, pp. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28

[25] I. Goodfellow *et al.*, "Generative adversarial networks," *Communications of the ACM*, vol. 63, no. 11, pp. 139–144, 2020. https://doi.org/10.1145/3422622

[26] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018, pp. 4510–4520. https://doi.org/10.1109/CVPR.2018.00474

[27] G. Hinton, O. Vinyals, and J. Dean, "Distilling the knowledge in a neural network," *arXiv preprint arXiv:1503.02531*, 2015.

[28] C. A. Tavera Romero, P. del R. Segovia de Maya, M. F. Díaz Velásquez, and D. F. Pérez Carvajal, "Development and impact of a mobile application that allows users to track their location on an educational institution campus," *Int. J. Interact. Mob. Technol.*, vol. 18, no. 1, pp. 110–132, 2024. https://doi.org/10.3991/ijim.v18i01.42905

[29] P. Sharma and A. K. Sinha, "A review on image encryption techniques for multimedia security," *Multimedia Tools and Applications*, vol. 79, no. 43, pp. 32229–32258, 2020.

[30] C. Blundo, A. De Bonis, and A. De Santis, "Improved schemes for visual cryptography," *Designs, Codes and Cryptography*, vol. 24, no. 3, pp. 255–278, 2001. https://doi.org/10.1023/A:1011271120274

[31] X. Wu and W. Sun, "Random grid-based visual secret sharing with abilities of OR and XOR decryptions," *Journal of Visual Communication and Image Representation*, vol. 24, no. 1, pp. 48–62, 2013. https://doi.org/10.1016/j.jvcir.2012.11.001

[32] M. Usman, I. Ahmed, M. I. Aslam, S. Khan, and U. A. Shah, "SIT: A lightweight encryption algorithm for secure Internet of Things," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 1, pp. 402–411, 2017. https://doi.org/10.14569/IJACSA.2017.080151

[33] R. R. Rakotomalala, P.-A. Guidec, C. Ménard, and J. Tisseau, "Adaptive bitrate streaming for mobile learning environments: A quality-of-experience driven framework," *IEEE Access*, vol. 10, pp. 74821–74838, 2022.

[34] A. Tao, "The dynamics of community engagement in distance education: A sociological analysis based on online learning platforms," *Int. J. Interact. Mob. Technol.*, vol. 18, no. 7, pp. 4–18, 2024. https://doi.org/10.3991/ijim.v18i07.48599

8 AUTHORS

Ms. S. Shibani Raju is a faculty member and research scholar at Directorate of Learning and Development, Department of Computing Technologies, SRM Institute of Science and Technology DLD. She is actively involved in teaching and research, with interests in signal processing, electronics, and emerging communication technologies. She contributes to academic excellence through student mentoring, curriculum development, and participation in scholarly activities such as conferences and publications (E-mail: shibanir@srmist.edu.in).

Dr. R. Thenmozhi is an Associate Professor at the Department of Computing Technologies, SRM Institute of Science and Technology, specialising in Computer Science and Software Engineering. With over 15 years of academic experience, her research focuses on networking, deep learning, and big data analytics. She has contributed to multiple publications, including journal articles, books, and conference proceedings, and has received over 100 citations with a notable h-index (E-mail: thenmozr@srmist.edu.in).