IMAGE WATERMARKING WITH MEMORY ACROSS REISSUANCE

Anonymous authors
Paper under double-blind review

ABSTRACT

Publishers often revise and reissue images, yet valid watermarks may persist on superseded releases or unapproved edits. Even a current release can contain regions last authorized at different revisions. We introduce and formalize *reissuable provenance*: determining an image's current, superseded, or unissued status and recovering each region's last authorized revision (its generation), with abstention when evidence is insufficient. *Generational Stratigraphy Fields* (GSF) combine watermark evidence with an authenticated live reference to the current release, enabling replacement or retirement without rewriting circulated copies. Each reissue freshly embeds new generation labels in approved regions and inherited labels elsewhere, preserving history without requiring old watermarks to survive successive edits. Our complete-output reissuance evaluation follows 480 histories through four releases, retaining each AI editor's entire output. GSF achieves 99.67% accuracy in classifying issued images as current or superseded and 94.09% recall of unissued edits. After three reissues, one generation map provides time and location: it recovers regional labels with 93.13% median accuracy and localizes the latest approved revision with 92.86% IoU. Regional labels remain readable under tested delivery transformations that impair issued-state recognition. Watermarks can thus express who issued an image, which release is authoritative now, and when each region was last authorized.

1 INTRODUCTION

Watermark verification answers whether an image carries a publisher's mark, and a verified image may be treated as authentic (Gowal et al., 2025; Nemecek et al., 2026; Coalition for Content Provenance and Authenticity, 2026a). Publication, however, is not always a one-time event. A publisher may revise and reissue an image to protect privacy, address safety risks, correct errors, or meet compliance requirements. We call each authorized release an *issue*. Others may edit an issued image without the publisher's approval, yet enough of the mark may survive for the result to verify (Bui et al., 2025; Lu et al., 2025; Nemecek et al., 2026). The superseded issue, the current issue, and the unissued edit therefore all point to the same trusted source while demanding different decisions. The publisher cannot simply clean up, either: once copies have spread across websites, group chats, and caches, correction is not recall; the publisher can replace or withdraw the image but cannot rewrite circulated copies (Melley, 2024; Gowal et al., 2025), whose marks may remain valid (Figure 1a). Nothing has failed: the mark is telling the truth about the past. But the verifier now needs an answer about the present. Is this the current issue, a genuine but superseded one, an edit that was never issued, or an image without sufficient evidence?

This is only half of the question: whole-image currency alone is insufficient because an issue need not change every region. The sky of a current issue may date from the first issue, the subject from the second, and the caption from the third; a version number says when the image was issued, not when each region entered authorized history. This distinction matters for auditing (Moreira et al., 2018; Nian et al., 2026). A disputed object or background change introduced at one revision persists through later, otherwise normal revisions (Figure 1b); to retrieve the record that approved it, an auditor needs each region's last authorizing revision, not a change mask for the latest edit. Figure 1 shows both halves: authentic does not mean current, and a disputed region keeps its origin through later issues. Both halves are becoming routine: instruction-guided AI editors put a revision one prompt away (Brooks et al., 2023; Black Forest Labs et al., 2025), and return a complete re-

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

rendering rather than a patch (Meng et al., 2022; Gu et al., 2026; Qi et al., 2026), so no part of the previous mark can be assumed to survive.

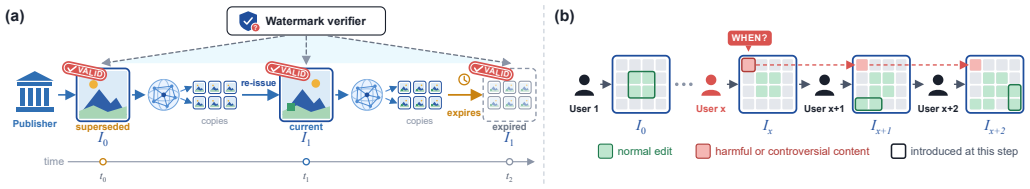

Figure 1: **Why present status and regional history are both necessary.** (a) **Authentic does not mean current.** A publisher issues I_0, corrects and reissues I_1, and I_1 later expires. The superseded issue, the current issue, and the retired copies all carry valid marks and receive the same VALID answer from a static verifier. (b) **A disputed edit propagates.** A harmful or controversial local change introduced at revision x persists through otherwise normal later revisions.

We call this joint task *reissuable provenance*. Its inputs are a query image, the issuer's key, and a small trusted reference to the lineage's current authorization state. A *lineage* is the set of issues descending from one source. Its first output is the query's present status: CURRENT for the active issue; SUPERSEDED for an issue displaced by a successor or explicitly retired; OPEN for an edit of the lineage that was never issued—a fact of its history, not a verdict; and UNKNOWN when the evidence is insufficient—an abstention, since a destroyed mark is not by itself an accusation. Its second output is a *field*: for each region, the authorized revision that last changed it. The two are coupled: one reissuance both moves the image's status and labels the approved regions. The two answers belong in different places. Present authority can change while the image does not, so no fixed image, key, and verifier can determine it from the pixels alone; it lives in the trusted reference, one entry per lineage. Regional history is fixed at issuance and travels inside every circulated copy: pixels persist when metadata is stripped, and any party entrusted with the issuer's key and the lineage's list of issues reads the history from the copy, without the trusted reference or the issuer's edit records. No existing watermark answers either half, for a structural reason.

Existing watermark pipelines pose one round (embed, edit, detect), so time and authorization never enter. A robust mark is expected to survive (Bui et al., 2025; Lu et al., 2025); a fragile mark is expected to break at the edit (Wong & Memon, 2001; Wu et al., 2023; Zhang et al., 2024; 2025); re-embedding is an overwrite attack (Bulychev et al., 2026) or a removal problem (Bui et al., 2025; Liu et al., 2025); and an editor's output is treated as adversarial, never as a draft. Each reports whether and where a mark is intact at one moment; none can say whether an issue is still current (an answer that changes after the image is fixed) or which approval a region came from. Common localization benchmarks hide the gap by restoring the premise AI editors removed, an unchanged protected background: localization is scored by inpainting a region and pasting it back into the protected image, so the background and its old mark survive by construction (Zhang et al., 2024; 2025), and supersession, approval, and inheritance never occur. Manifest standards such as C2PA do record revisions and revocation in signed metadata, which a watermark can help rediscover (Coalition for Content Provenance and Authenticity, 2026b;a), but the record describes one asset's history, not which issue of a lineage is current or which approval each region came from.

Where a one-round pipeline sees a credential that survives or breaks, our design reads the mark as evidence at three moments (when an image becomes OPEN, CURRENT, or SUPERSEDED), so that time and authorization enter. After an edit (OPEN), surviving source evidence ties the output to its lineage and a disrupted field marks an unresolved change: a broken mark is not the end of provenance, but the edit remains unauthorized until the issuer approves it. At approval (CURRENT), the issuer writes the successor a fresh, complete field—the new revision on the approved regions, inherited labels elsewhere—so continuity comes from reissuance, not survival: no predecessor mark needs to survive beyond the next reissuance. A region approved twice carries the later approval; the earlier approval remains in that issue's record. The same edit is untrusted before issuance and official after it; what separates the two is authorization, not pixels. After a later reissue (SUPERSEDED), the predecessor remains unchanged but is no longer current. We therefore keep this mutable fact in the trusted reference introduced above, an authenticated *live head*, like OCSP for certificates (Santesson et al., 2013). The publisher updates it without touching any circulated copy. Everything inheritable is written into each reissued image; only what cannot be frozen lives in the head.

Generational Stratigraphy Fields (GSF) realize this division. The field records each region's last authorized revision (its *generation*) as a keyed two-bit fingerprint, so its rate does not grow with the history; the price is that fingerprints of different generations collide in some regions as the history deepens. A spatial reader resolves such collisions using unambiguous neighboring regions and the coherence of revisions, which tend to occupy connected areas. The head says which issue, if any, is current; the field says which issue last shaped each region.

A lifecycle question needs a lifecycle evaluation. We retain the complete outputs of AI editors, in which essentially every block changes, and follow 480 histories of four issued generations through review, authorization, reissuance, branching, delivery, and later queries, scoring both outputs on the same trajectories (Section 5.1). Three findings follow. Supersession works without rewriting copies: GSF classifies issued images as current or superseded with 99.67% accuracy, and recalls 94.09% of unissued edits as OPEN, so surviving marks do not make drafts official. Spatial reading is necessary: on cells with ambiguous two-bit fingerprints, independent decoding accuracy falls to 25.6% at the deepest tested history while the spatial reader stays near 89%. One field answers when and where: it recovers each region's last authorizing revision with 93.13% accuracy, and its newest slice overlaps the approved revision with 92.86% IoU. Continuity thus survives legitimate revision and unresolved modification when survival, disruption, authorization, and reissuance are treated as moments of one lifecycle. Our contributions are:

- **A reissuance problem.** We formalize reissuable provenance as a four-way status decision coupled with each region's last authorized revision. Present authority is stored in a trusted mutable reference because no static in-image credential can express it.
- **Generational Stratigraphy Fields (GSF)**. We introduce a reissuance protocol whose authenticated live head exposes replacement and retirement without rewriting circulated copies. In its field instantiation, a constant-rate keyed field carries complete inherited regional history, and a candidate-equivariant spatial reader recovers that compressed history.
- **A complete-output reissuance evaluation.** We evaluate the joint task in realistic AI-editing workflows, retaining full outputs through review, authorization, reissuance, branching, delivery transformations, and later verification. The results show that GSF sustains issue-level authority and regional history together.

2 RELATED WORK

Prior work supplies the ingredients of reissuable provenance separately but does not natively couple live authorization state with in-image regional revision history.

Re-watermarking. Repeated watermarking predates this work (Mintzer & Braudaway, 1999): Multiple Re-Watermarking traces trading paths through accumulated receiver marks, WaterUp updates a distribution-stage identity, and TrustMark's remover supports removal and re-embedding (Mascher-Kampfer et al., 2006; Liu et al., 2025; Bui et al., 2025); a later mark can suppress an earlier one (Bulychev et al., 2026). Each embedding changes the message an image carries; these operations do not themselves define lineage-wide currentness or an inherited field of last-authorized regional revisions. GSF re-purposes the same interaction as a controlled revision within one lineage: an authorized reissue, through the live head, supersedes every unchanged predecessor (pixels unchanged, status changed), while an authenticated head update can retire the lineage. Each successor receives a freshly written field of last-authorized regional revisions, including inherited labels.

Manifest provenance. C2PA records provenance in signed, updatable manifests, checks signer revocation through OCSP, and rediscovers stripped manifests through soft bindings (Coalition for Content Provenance and Authenticity, 2026b;a; Collomosse & Parsons, 2024); it already has temporal records and mutable resolution. Its validation, however, remains asset-scoped: an active manifest establishes validity for the queried asset, not which issue of a lineage is currently authorized, while revocation concerns the signer's credential rather than one issue's authority. An application could assemble a CURRENT/SUPERSEDED/OPEN/UNKNOWN policy from repository and lineage state and replay regions along the selected branch for each region's last authorized revision, but neither is native and both live outside the pixels. GSF instead separates the placement: regional history trav-

els inside every reissued image, while an authenticated live head supplies the mutable information needed for lineage currentness, naming the current issue or declaring the lineage inactive (Section 3).

Spatial and forensic watermarks. Spatial watermarks recover localized or multiple messages (Sander et al., 2025; Hu et al., 2025), and forensic watermarks pair global attribution with local tamper evidence (Zhang et al., 2024; 2025; Chen et al., 2025; Yang et al., 2025; Gan et al., 2026; Liu et al., 2026). Both output a snapshot: where a message remains readable or tampering left evidence. Their native snapshot outputs do not specify how authorization state and regional revision labels evolve through reissuance. GSF reads the same local evidence in time: in an unissued edit, surviving and disrupted evidence supports recognition of an unresolved transition; after reissue, the generation map records the last authorized revision of each region. Each slice marks regions by their last authorized revision; the newest localizes the issue's latest revision. One field thus provides time and location without a separate localization output.

Robustness and evaluation under generative editing. Robust watermarks are tested under regeneration (Zhao et al., 2024; An et al., 2024) and instruction-guided editing (Hu et al., 2024; Lu et al., 2025), and two concurrent analyses model diffusion editing as an information-contracting channel (Gu et al., 2026; Qi et al., 2026); both focus on whether a static payload remains recoverable after editing. GSF asks instead what preserved and disrupted evidence means before and after issuer commitment, for generative and conventional edits alike. Evaluation also depends on which image is retained: paste-back preserves the carrier outside the edited region, whereas a complete output may disrupt it throughout. Forensic schemes include complete generative edits (Zhang et al., 2025; Gan et al., 2026; Liu et al., 2026), but evaluate post-edit recovery or tamper evidence rather than authorization transitions through reissuance. Our evaluation carries the complete editor output through review, reissuance, branching, and later queries; payload robustness thereby becomes lifecycle robustness.

3 Reissuable Provenance

This section formalizes the two outputs introduced above: a status relative to the live head and a field of last authorized revisions. Throughout, the issuer holds the watermark key K and an authenticated *registry* that records, for each lineage L, its issued labels $\mathcal{G}_L$ and its live head h.

Objects and authorization. Let $\mathcal{B} = \{B_i\}_{i=1}^{N}$ partition the image into cells. A lineage L has a finite set of issued labels $\mathcal{G}_L$ and a live head $h \in \mathcal{G}_L \cup \{\dagger\}$: $h \in \mathcal{G}_L$ names the current issue, and $h = \dagger$ declares none active. A proposed successor has a fresh label $g \notin \mathcal{G}_L$ and an issued parent $\mathrm{pa}(g) \in \mathcal{G}_L$; an editor produces an unissued output U from the parent's issued image. If the issuer approves U as the content X_g of a new issue, it also commits (records in the registry) an *authorized revision scope* $A_g \subseteq \{1, \dots, N\}$, the record of what it approves as revision g. Pixel difference, watermark disruption, or an authenticated edit record may propose the scope, but only the issuer's commitment makes it A_g. Authorization then updates the last-authorized-generation field

$$\tau_0(i) = 0, \qquad \tau_g(i) = \begin{cases} g, & i \in A_g, \\ \tau_{\mathrm{pa}(g)}(i), & i \notin A_g. \end{cases} \tag{1}$$

Cells in A_g advance to g; all others inherit the parent's last authorized revision, so the parent map covers branches as well as linear histories. The field τ_g is logical; Section 4.2 gives its encoding.

Verification semantics. The *event graph* of issuance and retirement events defines an image's true status; the verifier must infer it from evidence. A query Y yields lineage evidence $\lambda_L(Y) \in \{0, 1\}$, whether the evidence supports association with L, and issue evidence $\mu_L(Y) \in \mathcal{G}_L \cup \{\perp\}$, a supported match to an issued label, if any; the decoded lineage identifier selects L in the registry, which returns $\mathcal{G}_L$ and h. Both are evidence-supported conclusions, and the verifier's evidence-conditioned prediction is

$$\widehat{s}_L(Y; h) = \begin{cases} \mathsf{UNKNOWN}, & \lambda_L(Y) = 0, \\ \mathsf{CURRENT}, & \lambda_L(Y) = 1, \quad h \in \mathcal{G}_L,\ \mu_L(Y) = h, \\ \mathsf{SUPERSEDED}, & \lambda_L(Y) = 1, \quad \mu_L(Y) \in \mathcal{G}_L,\ \mu_L(Y) \neq h, \\ \mathsf{OPEN}, & \lambda_L(Y) = 1,\ \mu_L(Y) = \perp. \end{cases} \tag{2}$$

With $h = \dagger$, every matched issue is SUPERSEDED, whether displaced or retired, so updating h changes the meaning of an unchanged image. OPEN is predicted when lineage evidence remains but no issued-label match is supported; such an output may come from an editor or an attack, and the prediction authorizes nothing. The verifier returns $\mathcal{V}(Y; L, K, h) = (\widehat{s}, \widehat{\tau})$ with $\widehat{\tau}(i) \in \mathcal{G}_L \cup \{\perp\}$, where $\perp$ marks insufficient regional evidence. For an unissued output derived from parent r, the target field is τ_r: an unissued edit does not advance the field, and only issuance of g assigns g to A_g.

Necessity of the live head. Supersession and retirement change the status of an image that itself does not change. A verifier whose every trusted input is fixed at issuance (the image, its key, any static reference) receives identical inputs before and after such an event and returns the same answer, so it cannot move an image from CURRENT to SUPERSEDED; that status must come from a trusted input the issuer updates, and telling the $|\mathcal{G}_L|$ possible current issues and the inactive value apart takes at least $\lceil \log_2(|\mathcal{G}_L| + 1) \rceil$ bits of it (Appendix A).

4 METHOD

4.1 REISSUANCE PROTOCOL

Our reissuance protocol connects authorization, regional inheritance, and head-relative verification. For a realization supporting both task outputs, each issued image carries

$$\mathcal{P}_g = \big(P_L^{\mathrm{lin}}, P_{g,X_g}^{\mathrm{issue}}, C_g\big), \qquad \widetilde{X}_g = \mathrm{Embed}_K(X_g; \mathcal{P}_g), \tag{3}$$

where P_L^{lin} is lineage evidence, $P_{g,X_g}^{\mathrm{issue}}$ is issue evidence for label g and its issued content X_g, and C_g encodes the field τ_g. The encoding and embedding operator are implementation choices (Section 4.2). Lineage evidence may remain decodable when no issue match is supported; once an issued label is matched, the live head determines whether it is CURRENT or SUPERSEDED.

Editing and reissuance form one transition (Figure 2). An unissued output with lineage support but no supported issue match is predicted OPEN. If the issuer approves the output, it commits a revision scope A_g (Section 3). The issuer then applies Eq. (1), and the writer freshly embeds the complete payload $\mathcal{P}_g$ into the approved content X_g, including both new and inherited regional labels. Once the successor is embedded, the issuer atomically records g with its parent link in the registry and sets $h \leftarrow g$, making all earlier issues SUPERSEDED without rewriting their copies. Retirement sets $h \leftarrow \dagger$ without touching any image or payload; a later issuance reactivates the lineage.

4.2 GENERATIONAL STRATIGRAPHY FIELDS: A CONSTANT-RATE KEYED TWO-BIT ENCODING

GSF instantiates this protocol with a shared watermark carrier for the three terms in Eq. (3). Both global terms are keyed and error-corrected, with issue evidence additionally bound to a descriptor of the issued content (Appendix B). The field admits two encodings at the same two-bit rate: direct generation labels and keyed candidate fingerprints (Appendix K). Direct two-bit labels are collision-free for a fixed alphabet of at most four generations, but a label for an unbounded lineage needs $\lceil \log_2 |\mathcal{G}_L| \rceil$ bits per cell, so its rate grows with the history. GSF adopts keyed fingerprints for an extensible candidate set, trading local ambiguity for position-dependent compatibility and explicit no-match or ambiguous readouts. With a pseudorandom function F_K,

$$c_{L,a,i} = \mathrm{Trunc}_2\, F_K(L \,\|\, a \,\|\, i) \in \{0,1\}^2, \qquad C_g(i) = c_{L,\tau_g(i),i}. \tag{4}$$

The two bits are not a version number but a compatibility test against the registry's candidate set $\mathcal{G}_L$, so one interface serves any finite candidate set.

At constant rate, collisions are unavoidable: under an ideal PRF a wrong candidate matches the true two-bit value with probability $1/4$, so against d other candidates a cell identifies its generation uniquely with probability

$$p_{\mathrm{uniq}}^{(2)}(d) = \left(\frac{3}{4}\right)^d, \qquad p_{\mathrm{amb}}^{(2)}(d) = 1 - \left(\frac{3}{4}\right)^d. \tag{5}$$

Wider fingerprints collide less but, under a fixed slot budget, yield fewer and coarser cells. Under the same ideal-PRF model (keyed-iid) and a fixed slot budget, two bits maximize the worst-case

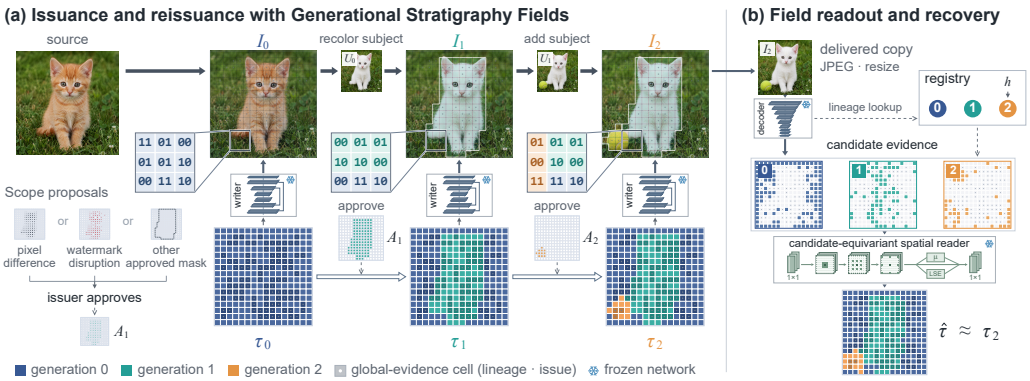

Figure 2: **Overview of reissuable provenance with GSF. (a)** Every issued image I_g carries its regional history as a 16×16 field τ_g. A reissue starts from an editor's unissued output U_g; the issuer approves a scope A_g (from pixel difference, watermark disruption, or another approved mask), and the writer rewrites the whole field: cells in A_g advance to generation g, the rest inherit theirs. Each cell holds a two-bit keyed fingerprint of its generation (insets: the same window across I_0, I_1, I_2); dotted cells carry lineage and issue evidence. **(b)** From a delivered copy, a frozen decoder reads the field bits and the lineage identifier; the registry supplies the candidate generations and the live head h, and the candidate-equivariant spatial reader resolves fingerprint collisions among the candidates to recover $\widehat{\tau}$.

expected number of uniquely identifying cells over $d \in \{1, 2, 3\}$, normalized by the best width at each depth; deeper histories gradually favor wider codes (Appendix L gives the law, its proof, and depth sensitivity).

At reissuance the writer embeds all three terms of Eq. (3) and rewrites every field cell: cells in A_g receive the fingerprint of g, all others the fingerprints of their inherited labels. Because $c_{L,a,i}$ is deterministic, a reissue refreshes the physical embedding without changing inherited values, so a successor carries its history itself rather than in residual pixels of its parent (embedding network and channel simulation: Appendix B).

4.3 Candidate-Equivariant Generation-Field Reader

Several generations can share a cell's fingerprint, so an independent two-bit decision cannot always identify the generation. The frozen decoder returns bit logits $\ell_i \in \mathbb{R}^2$ at cell i; for each issued generation $a \in \mathcal{G}_L$, $u_{i,a}$ sums the two bitwise log-likelihoods under $c_{L,a,i}$ (Eq. (14), Appendix F). A hard read z_i admits the candidates $\mathcal{C}_i(z_i) = \{a \in \mathcal{G}_L : c_{L,a,i} = z_i\}$: a singleton is an unambiguous temporal anchor, a larger set cannot be resolved from that cell alone, but neighboring cells can help because authorized revisions occupy coherent regions.

The reader forms a cell–candidate evidence field $Z_{i,a}$ from $u_{i,a}$, bitwise match evidence, and decoder reliability. A shared pointwise operator recalibrates likelihood, confidence, and margin evidence at each cell; a shared multiscale spatial operator then propagates reliable anchors across coherent revision regions. With Pool the concatenated mean and log-sum-exp over candidate planes,

$$E_a = \Phi_\psi^{\mathrm{sp}}\Big(\Phi_\psi^{\mathrm{pt}}(Z_a)\Big),$$
$$p_i(a) = \mathrm{softmax}_a\, \Psi_\psi(E_a, \mathrm{Pool}_{a' \in \mathcal{G}_L}(E_{a'}))_i\,,$$

where Φ_ψ^{pt}, Φ_ψ^{sp}, and Ψ_ψ share weights across candidate planes and Pool is permutation-symmetric. Hence $\mathbf{p}_\psi(\pi Z) = \pi \mathbf{p}_\psi(Z)$ for any permutation π of the candidate order (Eq. (15)), and the label map $\widehat{\tau}(i) = \arg\max_{a \in \mathcal{G}_L} p_i(a)$ carries no candidate-order semantics.

The reader returns $\widehat{\tau}$ where field evidence is readable and $\perp$ elsewhere; the predicted status still follows Eq. (2) from lineage and issue evidence.

5 EXPERIMENTS AND RESULTS

5.1 A COMPLETE-OUTPUT REISSUANCE EVALUATION

We follow 480 lineages, each rooted in one COCO (Lin et al., 2014) image, through four issues I_0–I_3. Between consecutive issues, one of four editors (FLUX.1 Kontext (Black Forest Labs et al., 2025), InstructPix2Pix (Brooks et al., 2023), Qwen Image Edit (Wu et al., 2025), SDXL-Turbo (Sauer et al., 2024)) returns a complete image, which the issuer approves and reissues; each root also keeps two outputs that are never issued, the editor's output after I_3 and one branch from I_1 or I_2 (2,400 editor calls). Every return is kept whole (the median fraction of field cells with any pixel change is 1.00 for all four editors), so the evaluation does not rely on preserved background pixels. Thresholded blockwise pixel difference proposes each scope, a median of about 150 of the 256 cells; the issuer commits it as A_g, the writer rewrites the whole field, and the head advances (Section 4). Here watermark disruption plays no part in the scope; it independently provides evidence to the verifier. True status is fixed by the event graph before verification; queries cover each issued image under its own head (4×480, CURRENT) and every later head (6×480, SUPERSEDED), and each unissued output under its parent's head and, for branches, later heads (3,129 queries, OPEN). Every query supplies the lineage L, its key K, and the head h, so the verifier decides among the four statuses for an image attributed to the lineage; negative queries, unmarked or of another lineage, are tested in Appendix N. One frozen verifier answers every query; metrics are status accuracy, OPEN recall, false-issue acceptance (an unissued output predicted CURRENT or SUPERSEDED), and for the field cell accuracy, mIoU over generations, and the IoU of the newest slice $\{i : \widehat{\tau}(i) = g\}$ with A_g; the root is the statistical unit. Construction details, the paste-back controls, and the statistical treatment are in Appendices C, I, and J.

5.2 GSF KEEPS AUTHORITY CURRENT ACROSS REISSUANCE

Queried as issued, status follows the live head without touching any issued image, and unissued outputs are read as OPEN (Table 1). Issued images are classified correctly in 99.67% of 4,800 queries, and no query reverses the head relation; of the unissued outputs, 94.09% are read as OPEN and 2.24% as an issue. Issue evidence separates the two: issued images keep a decodable issue match, while three quarters of unissued outputs lose it and a disrupted cluster catches most of the rest (Appendix D). Similar recall for outputs later reissued and never reissued (Table 1b) is consistent with a decision made on the commitment at query time rather than on subsequent reissuance.

Table 1: Lifecycle decisions (%). **(a)** Issued images under their own head (CURRENT) and under later heads (SUPERSEDED). **(b)** Unissued outputs. The 16 misclassified issued queries are read as OPEN (13) or UNKNOWN (3).

(a) Issued images: status accuracy (%) ↑

True status	$h=0$	$h=1$	$h=2$	$h=3$	Overall
CURRENT	99.58	99.79	99.79	98.96	**99.53**
SUPERSEDED	–	99.58	99.79	99.79	**99.76**

(b) Unissued outputs (%)

Output group	Queries	OPEN ↑	False issue ↓	UNKNOWN ↓
Later reissued	3×480	94.38	2.08	3.54
Never reissued	1,689	93.84	2.37	3.79
All	3,129	**94.09**	**2.24**	3.68

5.3 ONE GENERATION MAP PROVIDES TIME AND LOCATION

Once a revision is approved, the field records where and when it entered history. Repeated reissuance leaves a mosaic: 463 of 480 final fields carry all four generations, since cells outside each committed scope inherit their labels; a current issue thus holds more history than a whole-image version number can express. Table 2 reads that mosaic on the current issue at depths 1–3; at depth 3 the field reaches 93.13% map accuracy and 78.96% mIoU. mIoU falls from 93.87% at depth 1 because older slices become small and fragmented. On the 58.13% of cells whose fingerprint is ambiguous, the spatial reader reaches 88.27% accuracy against 25.56% for independent decoding. Figure 3 shows the audit path: a change that entered the lineage at generation a, whether an object, an occlusion, or a background, keeps label a while later committed scopes leave it untouched, so its region still points to the record that approved it. The newest slice of the same map overlaps the committed scope A_g with 92.86% IoU, so the field that dates each region also localizes the latest approved revision, without a separate localization output. Time and location are two readings of one field, and both hold on the same successors as Table 1.

Table 2: Regional history read from the current issue at depths 1–3 (%, medians over roots). Ambig.: share of the 160 field cells carrying generation fingerprints (the other 96 carry lineage and issue evidence) whose fingerprint is consistent with more than one issued generation; Cell/Spatial acc.: independent decoding and the spatial reader on those cells; Map acc./mIoU: all 160 cells; Newest IoU/F1: the newest slice against A_g.

	Ambiguous cells			Whole field and newest slice			
Depth	Ambig.	Cell acc. ↑	Spatial acc. ↑	Map acc. ↑	Map mIoU ↑	Newest IoU ↑	Newest F1 ↑
1	24.38	46.67	**88.64**	97.19	93.87	94.92	97.39
2	43.75	35.14	**89.23**	95.00	87.27	94.09	96.95
3	58.13	25.56	**88.27**	93.13	78.96	92.86	96.30

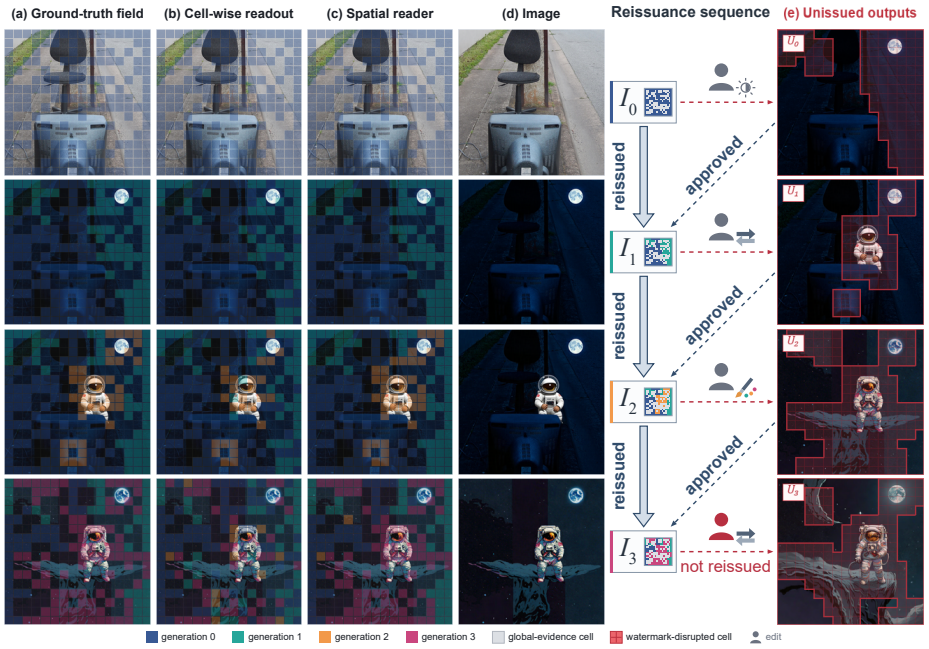

Figure 3: One field follows an authorized revision sequence $I_0 \rightarrow I_3$ and its unissued outputs. Colors give each region's last authorized generation, so an older change stays attributable after later revisions. Red cells mark disrupted field evidence in unissued outputs; gray cells carry lineage and issue evidence.

5.4 Status and Regional History under Delivery Transformations

Regional labels remain readable under tested transformations that impair status recognition. Applying one delivery transformation at a time to the final-head queries repeats the three status decisions and the map reading on the same current issue (Figure 4a–d). Under JPEG down to quality 75, downscaling to 256 px, blur up to $\sigma = 1.0$, and darkening, all three statuses stay above 93% without per-condition tuning, and OPEN recall does not fall. Brightening is the boundary for status: at $\times 1.15$ the current issue is still CURRENT in only 72.5% of cases, while the map read from the same images stays within 0.8 points of its untransformed value, as it does under every condition shown. Over the 31 transformations of Appendix E, map accuracy falls below 90% only under JPEG 30 and 192-px resizing, where at most 1% of issued images keep their status. Across the tested issued-image queries, status errors are OPEN or UNKNOWN rather than a CURRENT–SUPERSEDED reversal, and false acceptance of unissued outputs stays at or below 2.6%. Certifying an issue and reading its history rest on different evidence: a status needs a supported issue match for the whole image, which local field disruption or a failed issue-level check removes, whereas regional labels are read from the field itself and remain recoverable after that whole-image check fails.

5.5 The Reissuance Protocol Transfers across Watermark Carriers

The evaluated watermarks do not natively expose the combined live-head status, unissued-output, and regional-history interface. We therefore equip each with the reissuance protocol without

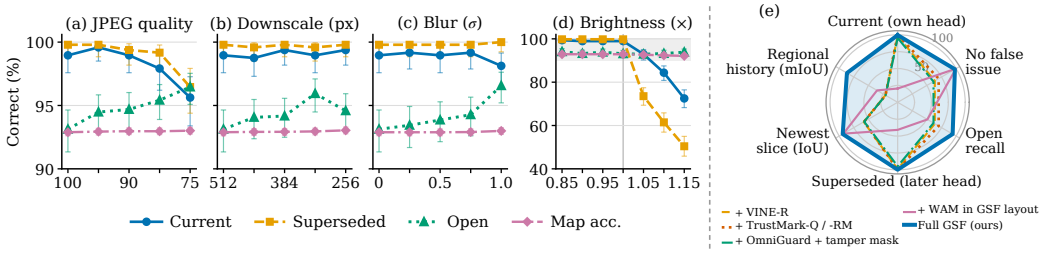

Figure 4: **(a–d)** Status decisions under ordinary delivery at head 3: CURRENT, the current issue I_3; SUPERSEDED, the initial issue I_0; OPEN, the never-issued output after I_3 and the branch, pooled; Map acc., generation-map accuracy of the same I_3 at depth 3, averaged over all 480 transformed images whatever their predicted status. **(e)** One polygon per system: CURRENT and SUPERSEDED accuracy, OPEN recall, the no-false-issue rate, and map mIoU and newest-slice IoU averaged over depths 1–3; TrustMark-RM coincides with TrustMark-Q.

retraining—a generation-conditioned issue message rewritten at every reissue, the live head, and the registry—plus OmniGuard's tamper mask and, for WAM, GSF's field layout (Appendix M). Table 3 thus tests support for the joint task, not the methods' native objectives. Every global-message watermark separates CURRENT from SUPERSEDED at 94.7–100%. This is a contribution of our protocol itself. Unissued recognition and inherited regional history, by contrast, need image evidence the protocol does not carry. The only such signal a global message carries is its destruction, which reads as OPEN, whereas a surviving message decodes to the old issue and is accepted. VINE-R and both TrustMark variants therefore recall 63.5–70.5% of unissued outputs and accept the remaining third as issues; OmniGuard reaches 61.3% recall, 26.3 points from its tamper mask, but its message survives more edits than any other, so its false-issue acceptance is the highest, 38.7%, against GSF's 94.0% and 2.1%. With the field layout, WAM localizes the newest revision as well as GSF (newest-slice IoU 93.44% against 92.86% at depth 3) but keeps little inherited history (24.30% mIoU against 78.96%): locating the latest change is not keeping the history behind it. GSF couples the live head with image evidence of disruption and history, and among the evaluated systems it alone closes all three parts of the task (Figure 4e).

Table 3: Evaluation on the same 480 roots and 2,400 edits (%). "+": a released watermark carrying our reissuance protocol. Issued columns: status accuracy under the own and later heads; unissued columns: the 6×480 unissued queries shared by all systems; mIoU: medians at depths 1–3. One rule for all systems: destroyed messages read as OPEN; carriers without a field never abstain (–); global messages are scored as fields labeling every cell with the head (gray). †WAM's generation block is readable on 442/480 roots; the rest count as errors.

	Issued images ↑		Unissued outputs			Generation-map mIoU ↑		
Existing carrier + our reissuance protocol	CURRENT	SUPERSEDED	OPEN ↑	False issue ↓	UNKNOWN ↓	d=1	d=2	d=3
+ VINE-R	98.2	98.9	63.5	36.5	–	28.91	19.40	14.70
+ TrustMark-Q	**99.9**	**100.0**	70.1	29.9	–	28.52	19.27	14.40
+ TrustMark-RM	99.6	99.8	70.5	29.5	–	28.52	19.21	14.65
+ OmniGuard + tamper mask	95.3	94.7	61.3	38.7	–	28.32	19.08	14.65
+ WAM in GSF layout†	20.4	40.8	50.8	0.4	48.8	49.05	32.38	24.30
Full GSF (ours)	99.5	99.8	**94.0**	2.1	3.9	**93.87**	**87.27**	**78.96**

6 CONCLUSION

We introduced reissuable provenance, a joint task beyond watermark validity: tracking current authorization, which changes with the live head, and each region's last authorized revision, updated only at issuance. The reissuance protocol connects authorization, reissuance, inheritance, and later verification, and transfers across watermark carriers. The GSF instantiation answers this task end to end: the live head and image evidence distinguish current issues, genuine but superseded issues, and unissued edits without rewriting circulated copies. Its generation map recovers inherited regional history rather than only locating the latest revision. Delivery experiments show that this history can remain readable when issue certification fails. The task, protocol, and complete-output evaluation provide a common basis for alternative codes, readers, and deployments.

ETHICS STATEMENT

This work involves no human subjects and collects no personal data. Source images are drawn from the public COCO dataset and used under its terms; the edited images are produced by released image editors, and the compared watermarks are released models, all used under their respective licenses, and no image is redistributed beyond what those terms permit. The reissuance protocol lets an issuer certify its own revisions; it judges authorization, not how content was produced, and authorizes nothing by itself: an OPEN decision marks an unresolved change and is not an accusation against whoever holds the image (Appendix E), and a SUPERSEDED decision changes only which issue is current. The live head is under the issuer's control and the watermark key is secret, so the mechanism confers no authority on third parties, and the paper reports the conditions under which its decisions fail or abstain (Appendices E and N). We foresee no other ethical concerns.

REPRODUCIBILITY STATEMENT

The task and verifier semantics are stated in Section 3 and formalized in Appendix A, which also proves the live-reference bound. The frozen instantiation is specified in Appendix B: payload coding and cell layout (Table 6), writer and decoder architectures (Tables 7 and 8), the training procedure and the physical-channel check (Table 9), blind alignment, and the evidence-routing constants used for every reported decision (Table 10). The evaluation is specified in Appendix C: root sampling, the prompt, seed, parent, and branch schedule fixed before generation (Tables 11 and 12), and the event-defined query schedule (Table 14). The statistical unit and interval procedure are given in Appendix J, whose Table 38 maps each claim to the experiment that supports it; the cross-method adapters and their reproducibility notes are in Appendix M. All results come from one frozen verifier without per-condition tuning. Code, the frozen checkpoints, the trajectory assignment, and every query row will be released upon publication.

AI USE STATEMENT

In this work, we used generative AI tools for the following tasks that require disclosure. Generating synthetic data sets: every edited image in the evaluation was produced by the released AI image editors named in Section 5.1 and Appendix C under a prompt, seed, parent, and branch schedule fixed before generation; these editors are the object of study, and no output was selected or regenerated by hand. Research ideation, conceptual framework, methodology, and experiment design: the core designs were provided by the authors; large language models assisted in refining their details and in writing the corresponding code, which the authors reviewed and tested. We have not used generative AI tools to formulate mathematical claims, to provide ingredients for or write proofs, or to interpret results; translation, dataset cleaning, and qualitative data analysis are not applicable to this work. Additionally, we used generative AI tools for recommended-disclosure tasks: drafting parts of the paper (portions of the appendices, later revised by the authors), editing the text for readability, identifying and summarizing related literature, and checking the consistency of tables and text against the experiment records. We have reviewed all AI-assisted work: every number, table, and claim was checked by the authors against the experiment records; model-drafted text was rewritten and verified by the authors; works suggested by the model were read by the authors before being cited; and model-assisted code was reviewed and its outputs checked against independent computations by the authors. We take responsibility for the final content of this work, including text, claims, and artifacts produced with the aid of generative AI.

REFERENCES

Bang An, Mucong Ding, Tahseen Rabbani, Aakriti Agrawal, Yuancheng Xu, Chenghao Deng, Sicheng Zhu, Abdirisak Mohamed, Yuxin Wen, Tom Goldstein, and Furong Huang. WAVES: Benchmarking the robustness of image watermarks. In *International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 1456–1492. PMLR, 2024.

Black Forest Labs, Stephen Batifol, Andreas Blattmann, Frederic Boesel, Saksham Consul, Cyril Diagne, Tim Dockhorn, Jack English, Zion English, Patrick Esser, Sumith Kulal, Kyle Lacey,

Yam Levi, Cheng Li, Dominik Lorenz, Jonas Müller, Dustin Podell, Robin Rombach, Harry Saini, Axel Sauer, and Luke Smith. FLUX.1 Kontext: Flow matching for in-context image generation and editing in latent space. *arXiv preprint arXiv:2506.15742*, 2025. doi: 10.48550/arXiv.2506.15742. URL https://arxiv.org/abs/2506.15742.

Raj Chandra Bose and Dwijendra Kumar Ray-Chaudhuri. On a class of error correcting binary group codes. *Information and Control*, 3(1):68–79, 1960. doi: 10.1016/S0019-9958(60)90287-4.

Tim Brooks, Aleksander Holynski, and Alexei A Efros. InstructPix2Pix: Learning to follow image editing instructions. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 18392–18402. IEEE, 2023.

Tu Bui, Shruti Agarwal, and John Collomosse. TrustMark: Robust watermarking and watermark removal for arbitrary resolution images. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 18629–18639, October 2025. URL https://openaccess.thecvf.com/content/ICCV2025/html/Bui_TrustMark_Robust_Watermarking_and_Watermark_Removal_for_Arbitrary_Resolution_Images_ICCV_2025_paper.html.

Maria Bulychev, Neil G. Marchant, and Benjamin I. P. Rubinstein. Watermarks attack watermarks: Re-watermarking as a generic removal strategy. *arXiv preprint arXiv:2605.16796*, 2026. doi: 10.48550/arXiv.2605.16796. URL https://arxiv.org/abs/2605.16796.

Yuzhuo Chen, Zehua Ma, Han Fang, Weiming Zhang, and Nenghai Yu. TAG-WM: Tamper-aware generative image watermarking via diffusion inversion sensitivity. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 16723–16732, 2025. URL https://openaccess.thecvf.com/content/ICCV2025/html/Chen_TAG-WM_Tamper-Aware_Generative_Image_Watermarking_via_Diffusion_Inversion_Sensitivity_ICCV_2025_paper.html.

C. J. Clopper and E. S. Pearson. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*, 26(4):404–413, 1934. doi: 10.1093/biomet/26.4.404.

Coalition for Content Provenance and Authenticity. C2PA soft binding API. Technical specification, 2026a. URL https://spec.c2pa.org/specifications/specifications/2.4/softbinding/Decoupled.html. Accessed 2026-08-19.

Coalition for Content Provenance and Authenticity. Content credentials: C2PA technical specification, version 2.4. Technical specification, 2026b. URL https://spec.c2pa.org/specifications/specifications/2.4/specs/C2PA_Specification.html. Accessed 2026-08-18.

John Collomosse and Andy Parsons. To authenticity, and beyond! Building safe and fair generative AI upon the three pillars of provenance. *IEEE Computer Graphics and Applications*, 44(3):82–90, 2024. doi: 10.1109/MCG.2024.3380168.

Bradley Efron. Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7(1): 1–26, 1979. doi: 10.1214/aos/1176344552.

Pierre Fernandez, Guillaume Couairon, Hervé Jégou, Matthijs Douze, and Teddy Furon. The stable signature: Rooting watermarks in latent diffusion models. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 22466–22477, 2023.

Zhenliang Gan, Chunya Liu, Yichao Tang, Binghao Wang, Shiwen Cui, Weiqiang Wang, and Xinpeng Zhang. GenPTW: Latent image watermarking for provenance tracing and tamper localization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 4085–4093, 2026. doi: 10.1609/aaai.v40i5.37412. URL https://ojs.aaai.org/index.php/AAAI/article/view/37412.

Sven Gowal, Rudy Bunel, Florian Stimberg, David Stutz, Guillermo Ortiz-Jimenez, Christina Kouridi, Mel Vecerik, Jamie Hayes, Sylvestre-Alvise Rebuffi, Paul Bernard, Chris Gamble, Miklós Z. Horváth, Fabian Kaczmarczyck, Alex Kaskasoli, Aleksandar Petrov, Ilia Shumailov, Meghana Thotakuri, Olivia Wiles, Jessica Yung, Zahra Ahmed, Victor Martin, Simon Rosen,

Christopher Savčak, Armin Senoner, Nidhi Vyas, and Pushmeet Kohli. SynthID-Image: Image watermarking at internet scale. *arXiv preprint arXiv:2510.09263*, 2025. doi: 10.48550/arXiv.2510.09263. URL https://arxiv.org/abs/2510.09263.

Fai Gu, Qiyu Tang, Te Wen, Emily Davis, and Finn Carter. When denoising becomes unsigning: Theoretical and empirical analysis of watermark fragility under diffusion-based image editing. *arXiv preprint arXiv:2603.04696*, 2026. doi: 10.48550/arXiv.2603.04696. URL https://arxiv.org/abs/2603.04696.

Dan Hendrycks and Kevin Gimpel. Gaussian error linear units (GELUs). *arXiv preprint arXiv:1606.08415*, 2016.

Alexis Hocquenghem. Codes correcteurs d'erreurs. *Chiffres*, 2:147–156, 1959.

Runyi Hu, Jie Zhang, Ting Xu, Jiwei Li, and Tianwei Zhang. Robust-Wide: Robust watermarking against instruction-driven image editing. In *European Conference on Computer Vision*, pp. 20–37. Springer, 2024. doi: 10.1007/978-3-031-72670-5_2.

Runyi Hu, Jie Zhang, Shiqian Zhao, Nils Lukas, Jiwei Li, Qing Guo, Han Qiu, and Tianwei Zhang. Mask image watermarking. In *Advances in Neural Information Processing Systems*, volume 38, 2025. doi: 10.52202/085713-4895. URL https://proceedings.neurips.cc/paper_files/paper/2025/hash/d74fb06615baaf6349007da7521223aa-Abstract-Conference.html.

Hugo Krawczyk, Mihir Bellare, and Ran Canetti. HMAC: Keyed-hashing for message authentication. RFC 2104, Internet Engineering Task Force, February 1997. URL https://www.rfc-editor.org/rfc/rfc2104.

Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C. Lawrence Zitnick. Microsoft COCO: Common objects in context. In *European Conference on Computer Vision*, pp. 740–755. Springer, 2014. doi: 10.1007/978-3-319-10602-1_48. URL https://link.springer.com/chapter/10.1007/978-3-319-10602-1_48.

Qingyu Liu, Yitao Zhang, Zhongjie Ba, Chao Shuai, Peng Cheng, Tianhang Zheng, and Zhibo Wang. Attack-resistant watermarking for AIGC image forensics via diffusion-based semantic deflection. In *International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=wyucYNGPiW.

Yanyan Liu, Bin Liu, Jie Zhang, Xiang Zhang, Zehua Ma, and Nenghai Yu. A watermark updating framework for multi-stage image content distribution. In *2025 IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1–6, Nantes, France, 2025. doi: 10.1109/ICME59968.2025.11209064. URL https://doi.org/10.1109/ICME59968.2025.11209064.

Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. In *International Conference on Learning Representations*, 2019.

Shilin Lu, Zihan Zhou, Jiayou Lu, Yuanzhi Zhu, and Adams Wai-Kin Kong. Robust watermarking using generative priors against image editing: From benchmarking to advances. In *International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=16O8GCm8Wn.

Andreas Mascher-Kampfer, Herbert Stögner, and Andreas Uhl. Multiple re-watermarking scenarios. In *Proceedings of the 13th International Conference on Systems, Signals, and Image Processing (IWSSIP)*, pp. 53–56, Budapest, Hungary, September 2006. URL https://www.cosy.sbg.ac.at/~uhl/multi.pdf.

Brian Melley. Why news organizations retracted the Princess of Wales photo. Associated Press report, republished by PBS NewsHour, March 2024. URL https://www.pbs.org/newshour/world/why-news-organizations-retracted-the-princess-of-wales-photo. Accessed 2026-09-02.

Chenlin Meng, Yutong He, Yang Song, Jiaming Song, Jiajun Wu, Jun-Yan Zhu, and Stefano Ermon. SDEdit: Guided image synthesis and editing with stochastic differential equations. In *International Conference on Learning Representations*, 2022.

Fausto Milletari, Nassir Navab, and Seyed-Ahmad Ahmadi. V-Net: Fully convolutional neural networks for volumetric medical image segmentation. In *International Conference on 3D Vision (3DV)*, pp. 565–571, 2016. doi: 10.1109/3DV.2016.79.

Fred Mintzer and Gordon W. Braudaway. If one watermark is good, are more better? In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 4, pp. 2067–2069, 1999. doi: 10.1109/ICASSP.1999.758338.

Daniel Moreira, Aparna Bharati, Joel Brogan, Allan Pinto, Michael Parowski, Kevin W. Bowyer, Patrick J. Flynn, Anderson Rocha, and Walter J. Scheirer. Image provenance analysis at scale. *IEEE Transactions on Image Processing*, 27(12):6109–6123, 2018. doi: 10.1109/TIP.2018.2865674.

Andreas Müller, Denis Lukovnikov, Jonas Thietke, Asja Fischer, and Erwin Quiring. Black-box forgery attacks on semantic watermarks for diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 20937–20946, 2025.

National Institute of Standards and Technology. Secure hash standard (SHS). FIPS PUB 180-4, U.S. Department of Commerce, August 2015.

Alexander Nemecek, Hengzhi He, Guang Cheng, and Erman Ayday. Authenticated contradictions from desynchronized provenance and watermarking. *arXiv preprint arXiv:2603.02378*, 2026. doi: 10.48550/arXiv.2603.02378. URL https://arxiv.org/abs/2603.02378.

Yi Nian, Aojie Yuan, Haiyue Zhang, Jiate Li, Li Li, Xiyang Hu, Hua Wei, Xiongye Xiao, Chaowei Xiao, and Yue Zhao. Auditable agents. *arXiv preprint arXiv:2604.05485*, 2026.

Dustin Podell, Zion English, Kyle Lacey, Andreas Blattmann, Tim Dockhorn, Jonas Müller, Joe Penna, and Robin Rombach. SDXL: Improving latent diffusion models for high-resolution image synthesis. In *International Conference on Learning Representations*, 2024.

Qian Qi, Jiangyun Tang, Jim Lee, Emily Davis, and Finn Carter. Editing away the evidence: Diffusion-based image manipulation and the failure modes of robust watermarking. *arXiv preprint arXiv:2603.12949*, 2026. doi: 10.48550/arXiv.2603.12949. URL https://arxiv.org/abs/2603.12949.

Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10684–10695, 2022.

Mehrdad Saberi, Vinu Sankar Sadasivan, Keivan Rezaei, Aounon Kumar, Atoosa Chegini, Wenxiao Wang, and Soheil Feizi. Robustness of AI-image detectors: Fundamental limits and practical attacks. In *International Conference on Learning Representations*, 2024.

Tom Sander, Pierre Fernandez, Alain Durmus, Teddy Furon, and Matthijs Douze. Watermark anything with localized messages. In *International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=IkZVDzdC8M.

Stefan Santesson, Michael Myers, Rich Ankney, Ambarish Malpani, Slava Galperin, and Carlisle Adams. X.509 internet public key infrastructure online certificate status protocol – OCSP. RFC 6960, Internet Engineering Task Force, June 2013. URL https://www.rfc-editor.org/rfc/rfc6960.

Axel Sauer, Dominik Lorenz, Andreas Blattmann, and Robin Rombach. Adversarial diffusion distillation. In *European Conference on Computer Vision*, pp. 87–103. Springer, 2024. doi: 10.1007/978-3-031-73016-0_6.

Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising diffusion implicit models. In *International Conference on Learning Representations*, 2021.

Yuxin Wen, John Kirchenbauer, Jonas Geiping, and Tom Goldstein. Tree-rings watermarks: Invisible fingerprints for diffusion images. In *Advances in Neural Information Processing Systems*, volume 36, pp. 58047–58063, 2023. doi: 10.52202/075280-2529.

Ping Wah Wong and Nasir Memon. Secret and public key image watermarking schemes for image authentication and ownership verification. *IEEE Transactions on Image Processing*, 10(10):1593–1601, 2001. doi: 10.1109/83.951543.

Chenfei Wu, Jiahao Li, Jingren Zhou, Junyang Lin, Kaiyuan Gao, Kun Yan, Sheng-ming Yin, Shuai Bai, Xiao Xu, Yilei Chen, Yuxiang Chen, Zecheng Tang, Zekai Zhang, Zhengyi Wang, An Yang, Bowen Yu, Chen Cheng, Dayiheng Liu, Deqing Li, Hang Zhang, Hao Meng, Hu Wei, Jingyuan Ni, Kai Chen, Kuan Cao, Liang Peng, Lin Qu, Minggang Wu, Peng Wang, Shuting Yu, Tingkun Wen, Wensen Feng, Xiaoxiao Xu, Yi Wang, Yichang Zhang, Yongqiang Zhu, Yujia Wu, Yuxuan Cai, and Zenan Liu. Qwen-Image technical report. *arXiv preprint arXiv:2508.02324*, 2025. doi: 10.48550/arXiv.2508.02324. URL https://arxiv.org/abs/2508.02324.

Xiaoshuai Wu, Xin Liao, and Bo Ou. SepMark: Deep separable watermarking for unified source tracing and deepfake detection. In *Proceedings of the 31st ACM International Conference on Multimedia*, pp. 1190–1201, 2023. doi: 10.1145/3581783.3612471.

Yuxin Wu and Kaiming He. Group normalization. In *European Conference on Computer Vision*, pp. 3–19. Springer, 2018. doi: 10.1007/978-3-030-01261-8_1.

Haoxin Yang, Bangzhen Liu, Xuemiao Xu, Cheng Xu, Yuyang Yu, Zikai Huang, Yi Wang, and Shengfeng He. StableGuard: Towards unified copyright protection and tamper localization in latent diffusion models. In *Advances in Neural Information Processing Systems*, volume 38, 2025. doi: 10.52202/085713-0495. URL https://proceedings.neurips.cc/paper_files/paper/2025/hash/15972a9575e0f03bf82f00aebeb40774-Abstract-Conference.html.

Zijin Yang, Kai Zeng, Kejiang Chen, Han Fang, Weiming Zhang, and Nenghai Yu. Gaussian shading: Provable performance-lossless image watermarking for diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 12162–12171, 2024.

Richard Zhang, Phillip Isola, Alexei A. Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 586–595, 2018.

Xuanyu Zhang, Runyi Li, Jiwen Yu, Youmin Xu, Weiqi Li, and Jian Zhang. EditGuard: Versatile image watermarking for tamper localization and copyright protection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 11964–11974, June 2024. URL https://openaccess.thecvf.com/content/CVPR2024/html/Zhang_EditGuard_Versatile_Image_Watermarking_for_Tamper_Localization_and_Copyright_Protection_CVPR_2024_paper.html.

Xuanyu Zhang, Zecheng Tang, Zhipei Xu, Runyi Li, Youmin Xu, Bin Chen, Feng Gao, and Jian Zhang. OmniGuard: Hybrid manipulation localization via augmented versatile deep image watermarking. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3008–3018, June 2025. URL https://openaccess.thecvf.com/content/CVPR2025/html/Zhang_OmniGuard_Hybrid_Manipulation_Localization_via_Augmented_Versatile_Deep_Image_Watermarking_CVPR_2025_paper.html.

Xuandong Zhao, Kexun Zhang, Zihao Su, Saastha Vasan, Ilya Grishchenko, Christopher Kruegel, Giovanni Vigna, Yu-Xiang Wang, and Lei Li. Invisible image watermarks are provably removable using generative AI. In *Advances in Neural Information Processing Systems*, volume 37, pp. 8643–8672, 2024. doi: 10.52202/079017-0276.

Jiren Zhu, Russell Kaplan, Justin Johnson, and Li Fei-Fei. HiDDeN: Hiding data with deep networks. In *European Conference on Computer Vision*, pp. 682–697. Springer, 2018. doi: 10.1007/978-3-030-01267-0_40.

A VERIFICATION SEMANTICS AND ATOMIC REISSUANCE

The central semantic change in reissuable provenance occurs without changing the queried bitmap. Once a successor becomes live, every earlier issue must be interpreted as SUPERSEDED; when the lineage is retired, no issue remains current. A small authenticated live head makes this new meaning available to every distributed copy at verification time, without rewriting the copy or its watermark. We formalize that behavior by separating event-defined truth, image evidence, and the verifier output supported by the evidence.

A.1 TRUTH, EVIDENCE, AND REGIONAL TARGETS

Let r denote an issued parent, g a proposed successor label, and $h \in \mathcal{G}_L \cup \{\dagger\}$ the live head, with $\dagger$ denoting an authenticated inactive lineage. Table 4 gives the complete distinction. For an unissued editor output, the edit event does not advance the regional field: the last-authorized-generation target remains τ_r wherever the field is readable and is $\perp$ where evidence is insufficient. Only an issuer commit assigns g to the committed scope.

Table 4: Event truth, observable evidence, and verifier outputs. Truth is defined by the authorization graph; predictions depend on evidence.

Query event	Event-defined truth	Supported issue evidence	Head role	Regional target
Issue $a = h \in \mathcal{G}_L$	CURRENT	matches a	selects a as live	τ_a
Issue $a \neq h$, $h \in \mathcal{G}_L$	SUPERSEDED	matches a	establishes $a \neq h$	τ_a
Any issue, $h = \dagger$	SUPERSEDED	matches its issue	declares no issue active	τ_a
Unissued output from r	OPEN	no committed issue match	none until issuance	τ_r or $\perp$
Insufficient/out-of-lineage evidence	UNKNOWN	unsupported	cannot be applied	$\perp$

These evidence channels have complementary roles. Lineage evidence supports association with L; issue evidence supports a match to one issue; the live head interprets that match as CURRENT or SUPERSEDED; physical disruption can expose an unresolved change; and issuer authentication alone turns a proposed spatial support into an authorized revision. Thus disruption may help predict OPEN, but it does not itself authorize content.

A.2 AUTHORIZATION AS AN ATOMIC REISSUANCE TRANSACTION

The implementation follows the following transaction, expressed without code-specific function names.

1. **Verify the parent.** Resolve an issued parent $r \in \mathcal{G}_L$ and interpret it relative to the current head h. The proposed label satisfies $g \notin \mathcal{G}_L$.
2. **Produce and review an unissued output.** An editor returns a complete image U. The event graph assigns $s^\star(U; h) = \mathsf{OPEN}$, while its regional last-authorized target remains τ_r.
3. **Propose a revision scope.** Physical image evidence or an authenticated edit record proposes R_g. The proposal has no authority by itself.
4. **Authorize and construct the successor.** The issuer accepts or revises the proposal to obtain A_g, applies Eq. (1), and embeds a complete issue payload $\mathcal{P}_g$ into the approved content. This step freshly embeds both labels introduced by g and labels inherited from r.
5. **Validate, record, and advance.** The completed successor is checked before visibility. The registry then atomically performs

$$\mathcal{G}_L \leftarrow \mathcal{G}_L \cup \{g\}, \qquad \mathrm{pa}(g) \leftarrow r, \qquad h \leftarrow g. \tag{6}$$

If construction or validation fails, Eq. (6) is not performed: g is not a legal candidate and U remains an unclosed OPEN event. This ordering prevents a partially written successor from becoming the live issue.

Retirement is a separate authenticated transaction: setting $h \leftarrow \dagger$ leaves the issue set, images, and embedded payloads unchanged while making every issue non-current. A later successful issuance transaction restores an active issue through $h \leftarrow g$.

A.3 Proof of the Live-Reference Necessity

Consider two worlds W_0 and W_1. In W_0, issued bitmap $\widetilde{X}_a$ is the live issue; in W_1, either an authorized successor has been committed or the lineage has been retired, so the correct status of the same $\widetilde{X}_a$ is SUPERSEDED. Hold fixed the bitmap, key, verifier program, immutable registry material, and every other trusted verifier input. A deterministic verifier receives the same input tuple in both worlds and must return the same answer. A randomized verifier likewise induces the same output distribution because its conditional input distribution is unchanged. Neither can return CURRENT in W_0 and SUPERSEDED in W_1 with certainty. Therefore at least one trusted input must vary when the authorization relation changes.

For a fixed issue set $\mathcal{G}_L$, the head must distinguish $|\mathcal{G}_L|$ possible active issues and one inactive value. A mutable coordinate with b bits has at most 2^b distinguishable states. Hence $2^b \geq |\mathcal{G}_L| + 1$ and $b \geq \lceil \log_2(|\mathcal{G}_L| + 1) \rceil$. This proves the lower bound stated in Section 3. A minimal mutable coordinate is therefore sufficient to give otherwise static image evidence a live authorization meaning.

A.4 Security Assumptions

The issuer controls the watermark key and an authenticated registry that provides integrity and freshness for h and the finite candidate set $\mathcal{G}_L$. Historical bitmaps may be copied, replayed, post-processed, or edited, but they cannot become a new issue unless the issuer authenticates the issuance transaction. A replayed predecessor is therefore SUPERSEDED, an unissued derivative remains OPEN, and evidence with no supported lineage association is UNKNOWN (Appendix N). The image experiments evaluate these lifecycle events and the delivery transformations specified in Appendices C and E.

A.5 Frozen Components and Traceability

Table 5 identifies the frozen components that determine the reported results. The anonymous artifact manifest records versions, producing scripts, and table/figure dependencies.

Table 5: Frozen components of the evaluation.

Artifact	Role
Writer–decoder checkpoint	physical two-bit channel
Carrier/verification core	layout, coding, synchronization
Protocol implementation	issue evidence and trajectory execution
Spatial-reader checkpoint	candidate-equivariant recovery
Trajectory assignment	roots, prompts, parents, seeds
Evaluation image set	roots and derived lifecycle images
Evaluation rows	status, robustness, reader, and maps
Localization controls	matched endpoint comparison
Editor repositories	complete-output generators

Together, the frozen components, event assignments, query rows, and image manifest connect every reported table to one fixed issuance graph.

B GSF Instantiation and Training Details

This section specifies the physical instantiation used throughout the evaluation. One shared writer–decoder carries lineage evidence, issue evidence, and a spatial generation field at a fixed operating point.

B.1 Canonical Grid and Payload Allocation

The verifier first represents an image on a canonical 512×512 canvas. The decoder produces a 16×16 grid, so one cell corresponds to a 32×32 pixel block and carries two binary logits plus a reliability logit. The 256 cells are partitioned into 32 lineage cells, 64 issue-evidence cells, and 160 generation-field cells (Table 6). The lineage and issue cells are sampled deterministically from the interior; the generation field contains the remaining 100 interior cells and 60 outer-ring cells. Thus,

after reserving global evidence for status verification, the same carrier still provides 160 history-bearing regions.

Table 6: Coding and layout of the frozen carrier. A two-bit generation entry is a keyed candidate fingerprint, not a binary generation number.

Component	Cells	Physical slots	Coding
Lineage evidence	32	64	36 bits $\rightarrow$ BCH$(63, 36, t{=}5)$
Issue evidence	64	128	64 bits $\rightarrow$ BCH$(127, 64, t{=}10)$
Generation field	160	320	two keyed bits per cell
Total	256	512	shared writer/decoder

The 36-bit lineage identifier is BCH encoded (Hocquenghem, 1959; Bose & Ray-Chaudhuri, 1960), XORed with an HMAC-SHA256 stream (Krawczyk et al., 1997; National Institute of Standards and Technology, 2015) in the lineage-key domain, and placed in 63 of the 64 available slots. Issue evidence begins with a low-frequency content descriptor: the issued image is converted to grayscale, resized to 128×128, partitioned into a 4×4 layout, and quantized to four bits per block. Its 64 bits are BCH encoded and keyed by the candidate issue label. At verification, every issued candidate is decoded and the correctable candidate with the fewest corrections is selected.

For a generation candidate a and field cell i, the implementation computes

$$c_{L,a,i} = \mathrm{LSB}_2\,\mathrm{HMAC\text{-}SHA256}(K, L \| \mathfrak{g}a \| i). \tag{7}$$

The operation is deterministic for fixed (K, L, a, i). Reissuance is fresh at the embedding level, not at the fingerprint-value level.

B.2 Logical Inheritance and Physical Full Rewrite

The logical update is local, but the physical write is complete. Given the committed target τ_g, the payload supplied to the writer is

$$C_g(i) = c_{L,\tau_g(i),i} \quad \text{for every field cell } i. \tag{8}$$

Cells inside A_g therefore carry fingerprints of g, while cells outside A_g freshly embed the deterministic fingerprints of their inherited labels. Issuance never relies on residual pixels from the predecessor and does not use masked physical writing. The optional synchronization template of the training code is disabled at issuance; alignment is described below.

B.3 Writer and Decoder Networks

Both networks are content conditioned and fully convolutional (Tables 7 and 8). A convolution–GroupNorm(8)–GELU block (Wu & He, 2018; Hendrycks & Gimpel, 2016) is used throughout unless stated otherwise. The writer receives RGB concatenated with a nearest-neighbor expansion of the two-bit field and predicts a bounded RGB residual.

Table 7: Frozen writer architecture. All convolutions use padding that preserves size unless stride is two.

Stage	Operation	Channels	Spatial size
Input	RGB + two-bit map	5	512^2
Stem	3×3, stride 1	$5 \rightarrow 64$	512^2
Down 1	3×3, stride 2	$64 \rightarrow 64$	256^2
Down 2	3×3, stride 2	$64 \rightarrow 128$	128^2
Middle	two 3×3 blocks	$128 \rightarrow 128$	128^2
Up 2	nearest upsample, skip, 3×3	$192 \rightarrow 64$	256^2
Up 1	nearest upsample, skip, 3×3	$128 \rightarrow 64$	512^2
Output	3×3, zero initialized, $\tanh$	$64 \rightarrow 3$	512^2

The predicted residual is multiplied by 0.1 and added to the normalized image before clipping. The writer contains 596,355 trainable parameters.

The decoder contains 2,329,155 trainable parameters. Candidate likelihoods use the two bit logits; the confidence logit and hard-bit margin enter the reader evidence tensor in Appendix F.

Table 8: Frozen decoder architecture. The final three channels are two bit logits and one confidence logit.

Stage	Operation	Channels	Spatial size
Stem	3×3, stride 1	$3 \to 64$	512^2
Down 1–2	two 3×3, stride 2	$64 \to 64 \to 128$	128^2
Down 3–5	three 3×3, stride 2	$128 \to 128 \to 256 \to 256$	16^2
Context	two 3×3 blocks	$256 \to 256$	16^2
Readout	1×1 convolution	$256 \to 3$	16^2

B.4 Writer–Decoder Training

We train for 100,000 optimizer steps with batch size 16 and AdamW (Loshchilov & Hutter, 2019) with learning rate 10^{-4}, default betas $(0.9, 0.999)$, default weight decay 0.01, and gradient-norm clipping at 1.0. Images were center cropped, resized to 512×512, and sampled from a 25,000-image COCO training directory disjoint from the 480 evaluation roots. The experiment assigns each image set a single role: these 25,000 images fit the writer and decoder; a separate 24-image COCO validation set checks the frozen physical channel (Table 9); development and calibration roots determine the reader and routing rules; and the 480 evaluation roots contribute only end-to-end scores. The 24-image set has zero image-ID overlap with the evaluation roots.

Let b denote the sampled payload and q the confidence target indicating whether both hard bits were correct at a cell. The training objective is

$$\begin{aligned}\mathcal{L} = \; & \mathrm{BCE}(D_{1:2}(T(\widetilde{X})), b) \\ & + 0.1\,\mathrm{BCE}(D_3(T(\widetilde{X})), q) \\ & + \lambda_q\big(100\,\mathrm{MSE}(\widetilde{X}, X) + \mathrm{LPIPS}(\widetilde{X}, X)\big).\end{aligned} \tag{9}$$

For the first 2,000 steps the channel is identity; over the next 10,000 steps the frozen distortion mixture is introduced, after which its six members are drawn uniformly. They are identity, real JPEG with quality uniformly sampled from 50–90 using a straight-through gradient, down/up resize with scale 0.5–1.0, Gaussian noise up to $2/255$, circular shifts of at most four pixels per axis, and a screenshot simulation (resize to 384, JPEG 70, resize to 512).

The quality weight starts at $\lambda_q = 0.05$. Every 100 steps, it is multiplied by 1.05 (capped at 200) if PSNR or LPIPS (Zhang et al., 2018) violates the quality target; otherwise, if BER exceeds 0.02, it is multiplied by 0.97 (floored at 0.05). The controller derives its quality targets from an independently measured reference-carrier reproduction under the same image normalization: the PSNR floor is the reference median minus 1 dB, and the LPIPS ceiling is twice the reference median. These rules yield PSNR ≥ 35.2439 dB and LPIPS ≤ 0.007310. The BER target of 0.02 governs the adaptive quality controller during training; status decisions use the frozen evidence-routing rules below. We retain the rule-derived precision so the training run can be reproduced exactly.

Table 9: Physical-channel characterization of the frozen checkpoint on 24 held-out COCO validation images, disjoint from both writer training and the 480 evaluation roots. BER is measured over the full grid before geometric support accounting or status routing.

Check	Observation	Protocol significance
Clean output	PSNR ↑: 45.646 dB; LPIPS ↓: 0.003431; median maximum error ↓: 13	Establishes the physical fidelity of the frozen carrier; the fidelity distribution of the 480 evaluation roots appears in Appendix J.
JPEG 75, resize to 384 and back, rotate 2°	BER ↓: 0.00098, 0.00000, and 0.00081	Confirms payload readout under representative delivery operators before status routing.
Screenshot simulation	BER ↓: 0.01571	Measures the harder composite delivery channel used in the training curriculum.
Crop 90% and resize	all-grid BER ↓: 0.07723	Exposes missing border support; the verifier accounts for that support before interpreting the payload.

Crop makes the distinction between raw bit error and status evidence explicit. A separate controlled support-accounting probe on development roots localized every crop error to the 60 border cells removed from view and measured zero BER on the common-view cells. The verifier records those border locations as predictable erasures before BCH decoding. The 0.07723 all-grid value therefore quantifies geometric support loss, while the status router consumes the erasure-aware evidence tensor. Appendices D and E measure the corresponding status behavior.

B.5 Blind Alignment and Frozen Evidence Routing

Alignment uses a deterministic search lattice and no reference image. Rotation and scale maximize the mean absolute decoder margin, a key-free score: the coarse search evaluates 13 rotations in $[-3^\circ, 3^\circ]$ and 11 isotropic scales in $[0.85, 1.15]$, and a second stage evaluates a 9×9 grid within $\pm 0.25^\circ$ and ± 0.015 scale of the coarse optimum. The verifier then searches integer translations with $(t_x, t_y) \in \{-16, -12, \ldots, 16\}^2$, scoring each candidate by agreement with the keyed expected signs of Eq. (7), and refines each coordinate within ± 3 pixels at unit steps. Cells brought in from outside the canvas are recorded as predictable erasures for BCH decoding. Evidence routing begins after canonicalization.

Threshold placement was completed on development data before the 480-root evaluation. Table 10 groups the resulting rules by evidence function and connects each placement rule to its end-to-end role. Together they form one frozen evidence router shared across transforms.

Table 10: Frozen evidence-routing rules. Development and calibration roots place every value; the resulting operating point remains fixed across all 480 evaluation roots, the clean images, and 31 delivery transformations.

Evidence family	Frozen rule	Placement rule (development data)	Protocol role
Cell readability	readable iff margin $> c_\star$, $c_\star = 4.2772$	Maximum of the per-root first-percentile margins over 12 development roots and five delivery transforms after predictable erasures (14,304 cell observations).	Produces the erasure mask consumed by downstream status inference.
Damage routing on clean queries	$G_{\mathrm{dead}} = 1, C_{\mathrm{cc}} = 1, B_{\mathrm{bg}} = 1/4$	The three coordinates were fixed jointly at one operating point on 144 clean queries from 12 development roots and 20 additional roots selected by a fixed hash rule. The status objective maximized event-defined OPEN while preserving status correctness on issued images and bounding false CURRENT and issued-image OPEN outcomes.	Separates localized disruption supporting OPEN from complete loss routed to UNKNOWN.
Unmatched field candidates	$C_{\mathrm{nokey}} = 1$	Smallest integer covering the benign calibration maximum while keeping the row-level false-trigger bound below 10^{-3}; confirmed on a separate 12-root calibration split.	More than one live cell matching no issued candidate supplies positive OPEN evidence.
Issue consistency	$\tau_{\mathrm{issue}} = 17/64$	99th percentile over 359 benign queries from 60 development roots, rounded up to the native $1/64$ distance grid.	Excess mismatch supplies positive OPEN evidence.
Current-issue safeguard	$\tau_{\mathrm{id}} = 1/64$ within $\lvert\theta\rvert < 0.5^\circ$, $\lvert s-1\rvert \le 0.001$	The pose predicate was fixed before the evaluation; the mismatch threshold is the rounded-up 99th percentile of 240 benign identity-pose queries on the same 60 roots.	Withholds CURRENT when current-issue evidence is insufficient, routing the internal rejection to UNKNOWN.

The router composes these signals in a fixed order. Low-margin cells first become erasures. Unmatched live candidates, undecodable issue evidence, a localized disrupted cluster, and issue inconsistency then provide positive OPEN evidence. Broad degradation without an issue match supports UNKNOWN. Once an issue is matched, equality with h yields CURRENT, whereas a different issued label yields SUPERSEDED. Appendix D evaluates this one operating point by status accuracy on issued images, OPEN recall, and false issue acceptance. Appendix E then applies it to the clean images and 31 delivery transformations without per-condition adjustment.

Static cost. The image-resident writer and decoder together contain 2.93M trainable parameters; the default spatial reader adds only 4,633. Registry and storage growth are reported in Appendix J.

C Complete-Output Reissuance Evaluation

We instantiate the reissuance protocol on COCO-derived sources and retain the complete RGB image returned by each editor. Every trajectory therefore passes through editing, review, authorization, reissuance, branching, and later head-relative verification. This design evaluates reissuable provenance across a realistic complete-output workflow rather than restoring generated regions to an unchanged watermarked background.

C.1 Roots, Trajectories, and Edit Events

We randomly sampled 480 images from the COCO dataset (Lin et al., 2014) as lineage roots. Each image was center cropped and resized to 512×512; every prompt, editor, seed, parent, and branch assignment was fixed before generation.

For each root we store

$$\text{source} \to I_0 \to U_0 \to I_1 \to U_1 \to I_2 \to U_2 \to I_3 \to U_3, \tag{10}$$

together with one unissued derivative U_p from either earlier issue I_1 or I_2 (Table 11). The issues are $I_0, \ldots, I_3$; $U_0, \ldots, U_3, U_p$ are complete editor returns prior to any commit. The first three main-chain outputs are later closed by successors. The final output and earlier-issue derivative remain unclosed, demonstrating that OPEN does not presuppose a future issue.

Table 11: Composition of the complete-output evaluation.

Object	Count
COCO-derived roots	480
Issues per root	4
Complete-output edits per root	5
Issue operations	4×480
Complete-output editor calls	2,400
Stored images per root	10
Homogeneous-editor chains	48 (12 per editor)
Mixed-editor chains	432
Branches from issue 1 / issue 2	249 / 231

The operation schedule contains 480 subject replacements, 168 recolorings, 120 material changes, 480 additions, 192 removals, 384 background replacements, 384 weather/lighting changes, and 192 global-style changes. Every root contains at least three operation classes and no immediate repeated class; removals are scheduled only when an added object exists. The frozen schedule was executed with the editor configurations of Table 12, without outcome-dependent regeneration, seed replacement, or sample deletion. Figure 5 shows three of the resulting lineages with the operation and editor of each reissue.

Table 12: Frozen complete-output editor configurations. All returned images were converted to RGB, resized to 512×512 with Lanczos when necessary, and stored losslessly as PNG.

Editor	Precision	Inference configuration	Edges
FLUX.1-Kontext-dev	bfloat16	28 steps, guidance 2.5	592
InstructPix2Pix	float16	30 steps, text guidance 8.0, image guidance 1.5	593
Qwen-Image-Edit-2509	bfloat16	fused Lightning LoRA, 8 steps, CFG 1.0	591
SDXL-Turbo img2img	float16	8 steps, guidance 0.0, strength 0.50	624

FLUX.1 Kontext, InstructPix2Pix, and Qwen-Image-Edit are instruction editors (Black Forest Labs et al., 2025; Brooks et al., 2023; Wu et al., 2025); SDXL-Turbo (Sauer et al., 2024), distilled from SDXL (Podell et al., 2024), is used as an image-to-image editor in the SDEdit manner (Meng et al., 2022). All 591 Qwen rows use the fused Lightning LoRA, eight steps, and CFG 1.0.

C.2 Complete Returns Change the Whole Carrier

For every edit we measured the mean absolute RGB change (8-bit units) and the fraction of 32×32 cells containing physical pixel change. Table 13 shows that the median changed-cell fraction is one for all four editors. This does not mean every change is semantically intended; it establishes that an unchanged-background assumption is not available when the complete return is retained.

C.3 From a Physical Proposal to an Authorized Scope

At reissuance, the issuer supplies a proposed revision scope and commits the region it approves. GSF accepts proposals derived from an edit mask, layer history, an authenticated operation record, pixel difference, or watermark disruption. For the complete-return experiment, we instantiate this input with deterministic *blockwise pixel difference* between the issued parent and editor return, computed before verification.

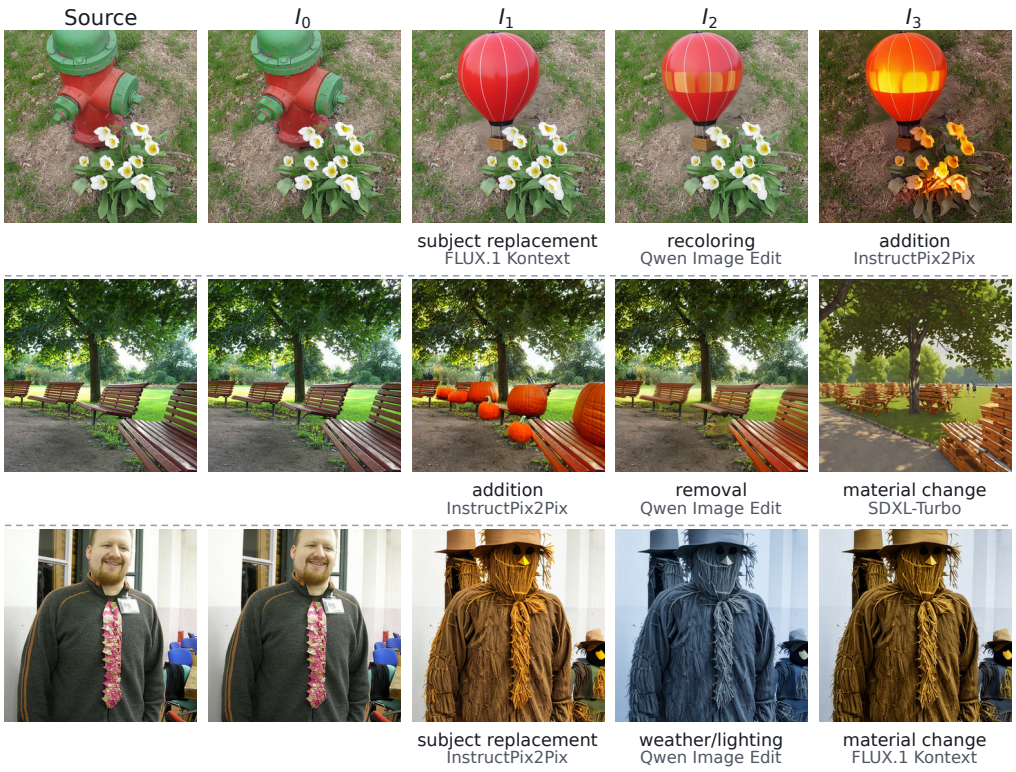

Figure 5: Three issued lineages from the evaluation: the source image, the initial issue I_0, and the reissues I_1–I_3, with the operation class and the editor of each reissue. Every reissue is a complete editor return, not a pasted patch, so the writer rewrites the whole field at each issuance.

Table 13: Physical change in retained complete outputs. Entries are computed from all 2,400 editor calls.

	Mean absolute RGB change			Changed-cell fraction		
Editor	p10	median	p90	minimum	p10	median
FLUX.1 Kontext	4.95	16.18	81.06	0.09	0.80	1.00
InstructPix2Pix	13.82	34.77	70.59	0.60	1.00	1.00
Qwen Image Edit	5.28	17.26	51.23	0.18	0.74	1.00
SDXL-Turbo	24.75	32.55	41.64	1.00	1.00	1.00

For an issued parent image I_r and complete editor return U, define

$$d(p) = \frac{1}{3}\sum_{c} |U_c(p) - I_{r,c}(p)|, \qquad \bar{d}_i = \frac{1}{|B_i|}\sum_{p \in B_i} d(p). \tag{11}$$

The initial proposal selects cells whose mean change exceeds a fixed relative and absolute threshold:

$$R_i^{(0)} = \mathbf{1}\left[\bar{d}_i > \max\{2, Q_{0.75}(\{\bar{d}_j\})\}\right]. \tag{12}$$

If fewer than 32 eligible interior generation cells are covered, the binary grid is dilated with a 3×3 kernel for at most eight iterations. This fixed construction supplies one reproducible proposed region for every complete return.

The issuer commits the approved region as $A_g \leftarrow R_g$, thereby updating τ_g and recording the successor. GSF then writes the committed scope into the successor and carries it forward as regional authorization history. A human-provided mask, layer record, or authenticated edit record enters the same commit-and-reissue path, allowing each deployment to choose proposal evidence suited to its authoring workflow.

The 256-cell logical grid defines the complete authorized revision field. In the evaluated carrier, 160 cells are configured as history-bearing observations, and generation-map metrics score all 160 of them; the remaining 96 cells provide lineage and issue evidence.

C.4 EVENT-DEFINED QUERY SCHEDULE

Table 14 enumerates all clean lifecycle queries. An issued bitmap can be queried at multiple later heads without being regenerated; this is precisely the test of head-relative semantics.

Table 14: Complete clean query schedule. Earlier-issue derivatives are evaluated at the heads for which their issued parent has the stated relation to the live issue.

Ground-truth class	Query condition	Query set
CURRENT	each issue under its own live head	4×480
SUPERSEDED	each predecessor under every later head	6×480
OPEN	pre-commit outputs later closed on the main chain	3×480
OPEN	unissued output from the terminal issue	1×480
OPEN	derivative from issue 1, queried at head 1	249
OPEN	derivative from issue 1, queried at head 2	249
OPEN	derivative from issue 2, queried at head 2	231
OPEN	all earlier-issue derivatives at final head 3	480

For an issued successor I_g, regional truth is τ_g. For an unissued output from r, global truth is OPEN regardless of decoder success, and the regional target remains τ_r. A regional output is scored only where field evidence is declared observable; elsewhere the verifier returns $\perp$.

C.5 METRICS AND STATISTICAL UNIT

Global results report per-query status accuracy or recall, false acceptance of an unissued output as an issue, and UNKNOWN abstention. Regional results report ambiguous-cell accuracy, all-field accuracy, mean IoU over generation labels present in the target, generation-specific IoU, and IoU/F1 of the newest-generation slice. The global and regional metrics in the joint table share the same successor and live head, directly pairing issue-level and regional-history evidence.

The root is the independent sampling unit. Heads, query conditions, cells, and post-processing settings are repeated measurements within a root. Tables state when an entry is a pooled count, a median over images, or a presence-conditioned IoU. Where bootstrap intervals (Efron, 1979) are used, roots are sampled with replacement and every within-root observation is kept together; cells are never treated as independent samples for uncertainty estimation.

D HOW REISSUANCE CLOSES THE TRANSITION

The complete result reveals a semantic transition that static watermark accuracy cannot express: an unchanged issue remains decodable while changing from CURRENT to SUPERSEDED, and a destroyed predecessor field becomes an unresolved OPEN transition until reissuance creates its successor. The tables below establish both halves on the complete query schedule. Their labels are assigned from the event graph before verification, so weak image evidence can cause an OPEN event to be missed but can never remove it from the denominator.

D.1 A LIVE HEAD MAKES STATIC ISSUES EXPIRE

Table 15 follows each unchanged issued bitmap as the live head advances. The same issue is CURRENT at its own head and SUPERSEDED at every later head; no image is regenerated for the later queries. This makes the experiment a direct test of head-relative semantics rather than a new decoding trial on a modified predecessor.

Across the ten issued-image query conditions, the combined status accuracy is 99.67%. More importantly for the supersession claim, the verifier never calls an issued predecessor CURRENT after the live head has advanced, and never calls the live issue SUPERSEDED. Its residual errors on issued images all withhold a status decision: 13 return OPEN and three return UNKNOWN. None reverses the current–superseded relation defined by the live head.

Table 15: Issued images under their own and later live heads. Entries are counts over 480 roots. "Open" and "Unk." are the two observed error destinations; there is no CURRENT↔SUPERSEDED confusion.

Ground truth	Query condition	Correct	Open	Unk.	Accuracy (%) ↑
CURRENT	issue 0 at head 0	478	2	0	99.58
	issue 1 at head 1	479	1	0	99.79
	issue 2 at head 2	479	1	0	99.79
	issue 3 at head 3	475	2	3	98.96
	all conditions under the issue's own head (4×480)	1911	6	3	99.53
SUPERSEDED	issue 0 at head 1	478	2	0	99.58
	issue 0 at head 2	479	1	0	99.79
	issue 1 at head 2	479	1	0	99.79
	issue 0 at head 3	479	1	0	99.79
	issue 1 at head 3	479	1	0	99.79
	issue 2 at head 3	479	1	0	99.79
	all conditions under a later head (6×480)	2873	7	0	99.76

D.2 REISSUANCE SEPARATES UNISSUED EDITS FROM ISSUES

The OPEN set includes both transitions that are later closed by a successor and outputs that remain unclosed. Table 16 reports every condition separately. The first three rows are the retained editor outputs immediately before issues 1–3 are committed. The final unissued output and derivatives from earlier issues have no assumed successor; their truth therefore cannot depend on a future reissue.

Table 16: Event-defined OPEN queries. "False issue" combines false CURRENT and false SUPERSEDED acceptance. Counts in each row sum to the stated query set.

Query condition	Query set	Open	False issue	Unk.	Open recall (%) ↑
pre-issue-1 output at head 0	480	461	9	10	96.04
pre-issue-2 output at head 1	480	452	11	17	94.17
pre-issue-3 output at head 2	480	446	10	24	92.92
pre-reissue editor outputs	3×480	1359	30	51	94.38
final unissued output at head 3	1×480	440	12	28	91.67
derivative from issue 1 at head 1	249	234	6	9	93.98
derivative from issue 1 at head 2	249	238	9	2	95.58
derivative from issue 2 at head 2	231	219	1	11	94.81
all earlier-issue derivatives at head 3	480	454	12	14	94.58
earlier-issue derivatives	1209	1145	28	36	94.71
all unissued outputs	3129	2944	70	115	**94.09**

Only 2.24% of unissued outputs are falsely accepted as an issue; 3.68% abstain as UNKNOWN. The earlier-issue results also show why an OPEN status cannot be defined as "the parent of an observed successor": the same semantics remain recoverable on final and earlier-issue outputs that are never committed.

The edit generators are not concentrated in one favorable chain type. The OPEN recall on homogeneous FLUX.1 Kontext, InstructPix2Pix, Qwen Image Edit, and SDXL-Turbo chains is 93.42% ($n = 76$), 97.47% ($n = 79$), 96.15% ($n = 78$), and 97.40% ($n = 77$), respectively. The much larger set of mixed-editor chains obtains 93.86% ($n = 2819$).

D.3 ISSUE EVIDENCE SEPARATES UNISSUED OUTPUTS FROM ISSUES

Issue evidence decodes on issued images and fails on edited ones: it matches an issue on 99.67% of issued queries, whereas on 74.94% of unissued outputs it does not decode (Table 18), and it keeps decoding under ordinary delivery (Appendix E). How the router reads that failure sets the operating point (Table 17). Abstaining on broad field loss forfeits most of it, at 14.29% recall; abstaining only on complete loss reads it as OPEN, 88.65%, at unchanged issued-image accuracy, so the recall gained comes from unissued outputs alone; counting any disrupted cluster as an unresolved change,

the operating point of Table 10, catches most outputs whose issue evidence survives an edit, reaching 94.09% and halving false issue acceptance for 0.21 points of issued-image accuracy.

Table 17: One routing chain at three operating points (%); the last is the point used throughout. The oracle is a diagnostic upper bound on the final unissued and earlier-issue subset (2×480), not a fourth operating point.

Routing rule	Issued acc. ↑	OPEN recall ↑	Issue FAR ↓
Abstain on broad field loss	**99.88**	14.29	4.92
Abstain only on complete field loss	**99.88**	88.65	4.92
+ read any disrupted cluster as OPEN	99.67	**94.09**	**2.24**
Oracle upper bound (subset)	–	98.23	–

Undecodable issue evidence or a disrupted cluster opens the transition; a decoded issue match is read against the live head as CURRENT or SUPERSEDED; reissuance then closes the transition with a new issue. The 98.23% oracle is a diagnostic upper bound on the stated 2×480 subset.

D.4 ABSTENTION SEPARATES MISSING EVIDENCE FROM AUTHORIZATION CLAIMS

Table 18 partitions all 3,129 unissued queries by their final decision and the dominant evidence route. It is a diagnostic decomposition, not a redefinition of truth. In particular, the 115 abstentions remain false negatives for OPEN recall.

Table 18: Dominant evidence routes for every event-defined unissued output.

Final prediction	Dominant evidence route	Count	Fraction (%)
OPEN	issue evidence does not decode	2345	74.94
	localized carrier-disruption support	543	17.35
	additional inconsistency evidence	56	1.79
UNKNOWN	insufficient or broadly degraded evidence	115	3.68
false issue	false CURRENT match	49	1.57
	false SUPERSEDED match	21	0.67

The dominant successful route is not literal pixelwise survival of the predecessor payload. It is the absence of a committed issue match and, where needed, localized carrier destruction. Fresh issuance reverses that condition by writing a complete successor payload and atomically advancing the live head. The evidence analysis therefore supports the central lifecycle mechanism: editing opens a transition, authorization closes it, and supersession changes the meaning of the otherwise unchanged predecessor.

E DELIVERY ROBUSTNESS: STATUS DECISIONS, REGIONAL HISTORY, AND THE BRIGHTENING BOUNDARY

The verifier carries three responsibilities through delivery. It must keep reading an issued image as CURRENT under its own head and as SUPERSEDED under a later head, it must keep reading an unissued output as OPEN, and it must keep the regional history of an issued image readable. Either of the first two alone is incomplete: a verifier that calls every damaged image OPEN obtains high OPEN recall while failing every issue it was meant to certify. This appendix therefore reports the complete matrices behind Figure 4 and Section 5.4: four endpoints and the generation map at depths 1–3, the clean images and 31 transformations, one frozen verifier and one frozen reader. It then documents the one boundary that matters for ordinary delivery, brightening, and the image property that explains why earlier issues cross it first.

E.1 POST-PROCESSING OPERATORS AND PAIRED ENDPOINTS

All inputs are stored and decoded as 8-bit RGB. Brightness multiplies RGB values followed by clipping; contrast scales values about the whole-image mean; gamma applies $255(x/255)^{\gamma}$; resizing is bicubic downsampling to the stated side length followed by bicubic restoration to 512×512; blur is Gaussian with the stated σ; JPEG uses the stated quality factor. The 31 transformations

are JPEG at quality 95, 90, 85, 75, 50, 40, and 30; double JPEG 75, which applies the quality-75 encode/decode twice; resizing to 448, 384, 320, 256, and 192 pixels; blur with σ of 0.25, 0.5, 0.75, and 1.0; brightness $\times 0.85$, 0.90, 0.95, 1.05, 1.10, 1.15, and 1.30; contrast $\times 0.80$, 1.25, and 1.50; gamma 0.80 and 1.25; and two composite platform proxies,

$$\begin{aligned} \text{Platform A}: &\quad \mathrm{JPEG}_{60}(\mathrm{Resize}_{384}(x)), \\ \text{Platform B}: &\quad \mathrm{JPEG}_{70}(\mathrm{Resize}_{256}(\mathrm{JPEG}_{80}(x))). \end{aligned}$$

These are deterministic image pipelines, not measurements from named commercial services. Figure 4(a–d) shows the JPEG, resize, blur, and brightness families up to JPEG 75, 256 pixels, $\sigma = 1.0$, and $\times 1.15$, with 95% Clopper–Pearson intervals (Clopper & Pearson, 1934) for the status lines and a bootstrap interval over roots for the map line; the tables below add the remaining levels and families.

Every condition is applied to four endpoints, all queried at head 3 and each with 480 roots: the current issue I_3 (truth CURRENT), the initial issue I_0 (truth SUPERSEDED), the never-issued output after I_3, written U_3, and the branch from I_1 or I_2, written U_p (both truth OPEN). These are the head-3 queries of Tables 15 and 16 under transformation, $4 \times 32 \times 480 = 61{,}440$ verifier queries in all. Every query uses the same verifier and the frozen rules of Table 10; there is no condition-specific calibration. The decisions were collected in two runs without new editor calls: one decoded I_3, U_3, and U_p under the clean images and the 21 coarser transformations; a later run added I_0 under those conditions and all four endpoints under the ten finer levels (JPEG 95, 90, and 85; brightness $\times 0.90$, 0.95, 1.05, and 1.10; blur 0.25 and 0.75; resize 448). The later run reproduced the earlier run's 1,920 clean decisions exactly, so the two runs are reported as one matrix.

Regional history under the same transformations. The generation map is read on the current issue at each depth, I_d at head d, as in Table 2, on the clean images and all 31 transformations: $480 \times 3 \times 32 = 46{,}080$ readings with the spatial reader (Appendix F). Post-processing does not change which generation wrote which cell, so the ground truth of every reading is the clean field of its root. The clean readings reproduce Table 2, and the head-3 status decisions of the same run reproduce Table 19 on all 15,360 current-issue queries. The same run also decided I_1 at head 1 and I_2 at head 2 under every condition, the own-head queries Table 19 lacks; Table 20 reports them beside the map. Reading the 46,080 maps took about 20 GPU-hours and no new editor call.

E.2 THE COMPLETE MATRIX

Table 19 reports all four endpoints. For the issued endpoints the truth is one issued status, so every OPEN or UNKNOWN entry is an error, not a conservative success; the two columns that would record a CURRENT–SUPERSEDED reversal are zero in every row and are omitted. For the unissued endpoints the pooled OPEN column is the OPEN line of Figure 4.

E.3 ENVELOPE AND BOUNDARY

Eighteen conditions keep all three statuses above 93%. The clean images; JPEG at quality 95, 90, 85, and 75; double JPEG 75; resize 448, 384, 320, and 256; blur at every tested σ; darkening to $\times 0.95$, 0.90, and 0.85; and contrast $\times 0.80$ all keep CURRENT retention at 95.21–99.58%, SUPERSEDED retention at 94.17–100.00%, and pooled OPEN recall at 93.13–96.77%. Moderate compression, resizing, blur, repeated JPEG encoding, darkening, and contrast reduction therefore do not force a choice between preserving an issue and recognizing an unissued output; the three-way distinction survives a non-pristine delivery path. The depth-3 map accuracy is at least its clean value of 93.13% in every one of these conditions (Table 20).

Fourteen conditions cross the boundary, and issued images fail toward OPEN or UNKNOWN. The boundary is sharpest under JPEG 30, resize 192, and Platform B, where CURRENT retention falls to 0.00–2.50% and SUPERSEDED retention to 0.00–0.63%, with 91.67–95.63% of current issues read as OPEN. JPEG crosses between quality 75 and 50 (95.63% to 35.00% CURRENT), resizing between 256 and 192 pixels (99.38% to 1.04%), and blur does not cross within $\sigma \leq 1.0$. Brightening and contrast gain cross more gradually: CURRENT retention under brightness $\times 1.05$, 1.10, 1.15, and 1.30 is 92.92%, 84.38%, 72.50%, and 39.38%, and SUPERSEDED retention of the initial issue 73.54%, 61.46%, 50.42%, and 24.58%. Gamma is the one family whose errors go to

Table 19: The complete delivery matrix at head 3 (% of 480 roots per endpoint). I_3 is CURRENT in truth and I_0 SUPERSEDED; the omitted columns, I_3 predicted SUPERSEDED and I_0 predicted CURRENT, are zero in every row. U_3 and U_p are OPEN in truth; the pooled columns combine them (960 queries), with false issue a prediction of CURRENT or SUPERSEDED. Unk. is UNKNOWN, including decisions withheld by the current-issue safeguard.

	Current issue I_3			Initial issue I_0			Open ↑		Pooled U_3 and U_p		
Condition	Current ↑	Open ↓	Unk. ↓	Superseded ↑	Open ↓	Unk. ↓	U_3	U_p	Open ↑	False issue ↓	Unk. ↓
Clean	98.96	0.42	0.63	99.79	0.21	0.00	91.67	94.58	93.13	2.50	4.38
JPEG 95	99.58	0.00	0.42	99.79	0.21	0.00	92.50	96.46	94.48	2.40	3.13
JPEG 90	98.96	0.42	0.63	99.38	0.63	0.00	93.13	96.25	94.69	2.08	3.23
JPEG 85	97.92	1.25	0.83	99.17	0.83	0.00	94.38	96.46	95.42	1.88	2.71
JPEG 75	95.63	3.96	0.42	96.46	3.54	0.00	95.00	97.92	96.46	1.77	1.77
JPEG 50	35.00	64.79	0.21	38.96	61.04	0.00	98.54	98.96	98.75	0.52	0.73
JPEG 40	1.67	94.79	3.54	4.17	94.79	1.04	98.96	99.79	99.38	0.00	0.63
JPEG 30	0.00	91.67	8.33	0.00	90.42	9.58	99.58	99.58	99.58	0.00	0.42
Double JPEG 75	95.21	3.96	0.83	94.17	5.83	0.00	95.42	98.13	96.77	1.77	1.46
Resize 448	98.75	0.63	0.63	99.58	0.42	0.00	92.29	95.83	94.06	2.60	3.33
Resize 384	99.38	0.21	0.42	99.79	0.21	0.00	91.88	96.46	94.17	1.98	3.85
Resize 320	98.96	0.63	0.42	99.58	0.42	0.00	93.96	97.92	95.94	1.56	2.50
Resize 256	99.38	0.21	0.42	99.79	0.21	0.00	92.71	96.46	94.58	1.35	4.06
Resize 192	1.04	94.17	4.79	0.42	92.29	7.29	96.88	98.33	97.60	0.00	2.40
Blur 0.25	99.17	0.21	0.63	99.79	0.21	0.00	92.08	94.79	93.44	2.50	4.06
Blur 0.5	98.96	0.42	0.63	99.79	0.21	0.00	91.67	96.04	93.85	2.29	3.85
Blur 0.75	99.17	0.42	0.42	99.79	0.21	0.00	92.71	95.83	94.27	2.29	3.44
Blur 1.0	98.13	0.83	1.04	100.00	0.00	0.00	95.63	97.50	96.56	1.25	2.19
Brightness ×0.85	99.38	0.42	0.21	99.79	0.21	0.00	92.92	95.21	94.06	2.29	3.65
Brightness ×0.90	98.96	0.42	0.63	99.79	0.21	0.00	91.25	95.21	93.23	2.29	4.48
Brightness ×0.95	98.96	0.42	0.63	99.79	0.21	0.00	91.88	95.83	93.85	2.40	3.75
Brightness ×1.05	92.92	6.04	1.04	73.54	26.46	0.00	90.83	95.00	92.92	2.40	4.69
Brightness ×1.10	84.38	13.33	2.29	61.46	38.54	0.00	91.25	95.21	93.23	2.19	4.58
Brightness ×1.15	72.50	21.88	5.63	50.42	49.58	0.00	91.88	95.83	93.85	1.77	4.38
Brightness ×1.30	39.38	50.42	10.21	24.58	75.21	0.21	94.17	97.29	95.73	0.94	3.33
Contrast ×0.80	99.17	0.21	0.63	99.79	0.21	0.00	91.46	95.42	93.44	2.40	4.17
Contrast ×1.25	64.58	29.79	5.63	31.04	68.96	0.00	91.04	95.63	93.33	1.46	5.21
Contrast ×1.50	31.25	61.04	7.71	10.42	89.38	0.21	94.58	95.83	95.21	0.94	3.85
Gamma 0.80	75.00	0.21	24.79	99.79	0.21	0.00	91.04	94.79	92.92	2.50	4.58
Gamma 1.25	66.67	3.33	30.00	93.96	6.04	0.00	92.08	95.83	93.96	1.77	4.27
Platform A	58.13	41.88	0.00	44.38	55.63	0.00	98.75	97.92	98.33	0.52	1.15
Platform B	2.50	95.63	1.88	0.63	96.46	2.92	98.96	99.58	99.27	0.00	0.73

UNKNOWN rather than OPEN: at gamma 0.80 and 1.25 the current issue is withheld in 24.79% and 30.00% of cases while the initial issue stays SUPERSEDED in 99.79% and 93.96% (Section E.4). In no condition does either issued endpoint acquire the other issued status: over the 30,720 issued queries of Table 19 there is no CURRENT–SUPERSEDED reversal. Broad degradation removes the positive issue match, so the verifier withholds the status it can no longer support; it never assigns the wrong one. Where the status crosses, the map does not: JPEG 50 leaves 35.00% of current issues CURRENT but 93.13% of their cells correctly labeled.

The map outlasts the status. Across the 31 transformations, the depth-3 map accuracy of the current issue falls below its clean value only where cells become unreadable: to 92.50% under Platform B, 91.88% under brightness ×1.30 and contrast ×1.50, 91.25% under JPEG 40, 87.50% under resize 192, and 85.00% under JPEG 30 (means 92.12%, 90.58%, 90.47%, 90.08%, 85.53%, and 82.89% against 92.89% clean). The largest losses at depths 1 and 2 fall under the same two conditions, JPEG 30 (93.13% and 88.75%) and resize 192 (92.50% and 90.00%), and the depth order of Table 2, mIoU falling from depth 1 to depth 3, holds in every condition. The status decision fails earlier because it reacts to evidence the map does not need. A localized cluster of unreadable cells is positive OPEN evidence (Appendix B): under JPEG 50 such a cluster appears in 65.0% of current issues while the median share of readable field cells is 96.9%, so 35.00% remain CURRENT and the map is unchanged; under gamma at least nine current issues in ten keep every cell readable, the current-issue safeguard withholds 24.79% and 30.00% (Section E.4), and the map is again unchanged. Where cells are lost, the map loses fewer labels than the field loses readable cells: at JPEG 30 the median readable share is 61.3% and 85.00% of cells are still labeled correctly, and the spatial reader's gain over independent decoding on ambiguous cells stays between 33.4 and 63.1 points in every condition and depth.

A surviving status carries the clean map. For images whose status survives a transformation, the map is the clean map: on the roots read as CURRENT, the paired change in map accuracy relative to the clean image has a median of 0.00 points at depths 1–3 in every condition with at least 30 such roots, and a mean within ± 0.25 points. A CURRENT decision under delivery is therefore accompanied by regional history as accurate as on the clean image.

Table 20: Regional history under delivery (medians over 480 roots, %). The current issue at each depth, I_d at head d, is transformed and read with the spatial reader: map accuracy and mIoU over the 160 field cells at depths 1–3 and the newest-slice F1 against A_3 at depth 3. The last three columns give, for the same queries, CURRENT retention of I_1, I_2, and I_3 at their own heads; the I_3 column repeats the delivery matrix. The clean row is the untransformed result.

	Map acc. ↑			Map mIoU ↑			Newest F1 ↑	CURRENT at own head ↑		
Condition	d=1	d=2	d=3	d=1	d=2	d=3	d=3	I_1	I_2	I_3
Clean	97.19	95.00	93.13	93.87	87.27	78.96	96.30	99.79	99.79	98.96
JPEG 95	97.50	95.00	93.13	93.84	87.11	79.18	96.30	99.58	99.79	99.58
JPEG 90	96.88	95.00	93.13	93.87	87.11	79.12	96.37	99.58	99.38	98.96
JPEG 85	96.88	95.00	93.13	93.85	87.00	79.03	96.34	98.75	97.92	97.92
JPEG 75	96.88	95.00	93.13	93.81	86.96	79.05	96.32	96.25	96.67	95.63
JPEG 50	96.88	95.00	93.13	93.62	86.58	78.32	96.06	41.88	36.25	35.00
JPEG 40	96.25	93.13	91.25	92.33	82.48	72.76	94.74	4.58	3.75	1.67
JPEG 30	93.13	88.75	85.00	86.40	73.78	61.45	90.25	0.00	0.21	0.00
Double JPEG 75	96.88	95.00	93.13	93.83	86.97	79.08	96.32	95.21	95.21	95.21
Resize 448	96.88	95.00	93.13	93.87	87.34	79.13	96.30	99.79	99.58	98.75
Resize 384	96.88	95.00	93.13	93.82	87.31	79.13	96.31	99.38	99.58	99.38
Resize 320	96.88	95.00	93.13	93.87	87.13	79.23	96.32	99.58	99.58	98.96
Resize 256	97.19	95.00	93.75	93.85	87.13	79.59	96.37	99.58	99.58	99.38
Resize 192	92.50	90.00	87.50	85.25	75.05	67.31	91.46	0.21	0.21	1.04
Blur 0.25	97.19	95.00	93.13	93.87	87.27	79.03	96.30	99.79	99.79	99.17
Blur 0.5	96.88	95.00	93.13	93.81	87.37	79.06	96.30	100.00	99.79	98.96
Blur 0.75	97.19	95.00	93.13	93.87	87.19	79.11	96.30	99.79	99.58	99.17
Blur 1.0	96.88	95.00	93.13	93.79	87.00	79.29	96.32	99.17	99.38	98.13
Brightness $\times 0.85$	97.19	95.00	93.13	93.87	87.27	79.09	96.30	99.79	99.58	99.38
Brightness $\times 0.90$	97.50	95.00	93.13	93.87	87.27	78.98	96.33	99.58	99.58	98.96
Brightness $\times 0.95$	97.50	95.00	93.13	93.87	87.27	79.03	96.30	99.79	99.79	98.96
Brightness $\times 1.05$	96.88	95.00	93.13	93.69	86.71	78.79	96.24	81.88	89.79	92.92
Brightness $\times 1.10$	96.88	95.00	93.13	93.56	86.40	78.28	96.08	69.38	78.13	84.38
Brightness $\times 1.15$	96.88	95.00	93.13	93.20	86.08	78.13	96.00	54.17	66.25	72.50
Brightness $\times 1.30$	96.25	93.75	91.88	92.14	83.84	75.86	95.51	25.42	32.50	39.38
Contrast $\times 0.80$	97.19	95.00	93.13	93.87	87.35	78.98	96.31	99.58	99.79	99.17
Contrast $\times 1.25$	96.88	94.38	93.13	93.15	85.59	77.72	95.96	42.71	56.46	64.58
Contrast $\times 1.50$	95.63	93.75	91.88	91.42	83.30	74.87	95.34	16.46	24.58	31.25
Gamma 0.80	97.50	95.00	93.13	93.83	87.18	78.96	96.30	74.17	71.25	75.00
Gamma 1.25	97.50	95.00	93.13	93.87	87.18	79.09	96.30	63.13	66.25	66.67
Platform A	96.88	95.00	93.13	93.72	86.67	79.23	96.30	61.88	56.25	58.13
Platform B	96.25	94.38	92.50	92.56	84.63	77.85	95.65	1.25	2.50	2.50

Unissued outputs stay OPEN throughout. Pooled OPEN recall never falls below 92.92% (the single-endpoint minimum is 90.83%, U_3 under brightness $\times 1.05$) and rises under heavy damage to 99.58% at JPEG 30, because damage adds disruption evidence. Pooled false acceptance of an unissued output as an issue stays at or below 2.60% (resize 448), and pooled UNKNOWN at or below 5.21% (contrast $\times 1.25$). Read together with the previous paragraph, the transformations never make a draft official; they only make a damaged issue lose its status.

The head changes the label, not the evidence. The initial issue was also queried at head 0, where its truth is CURRENT, under the clean images and the 21 coarser transformations in the earlier run. Pairing those decisions with the head-3 decisions of the same 10,560 images, 99.19% agree up to relabeling (CURRENT or withheld at head 0 against SUPERSEDED at head 3, OPEN against OPEN, UNKNOWN against UNKNOWN). The transformed-image results therefore hold for the head-relative decision as such, not for one head.

E.4 WHY EARLIER ISSUES FAIL FIRST UNDER BRIGHTENING

Under every brightening and contrast-gain level, retention falls with the age of the issue: at brightness $\times 1.15$ the initial issue keeps SUPERSEDED in 50.42% of cases, and I_1, I_2, and I_3 keep CURRENT at their own heads in 54.17%, 66.25%, and 72.50% (Table 20); at contrast $\times 1.25$ the four

figures are 31.04%, 42.71%, 56.46%, and 64.58%. This runs against the natural expectation that the least edited image is the most robust carrier. A post-hoc measurement of the untransformed images, Table 21, gives the reason. Brightening multiplies 8-bit values and clips at 255, so it destroys carrier evidence in exactly the pixels that were already bright, and every pass through a diffusion editor compresses the highlight tail: at nearly the same mean intensity, the 99th-percentile intensity is 242 for I_0 and 232, 228, and 222 for I_1, I_2, and I_3, with the unissued outputs U_3 and U_p in the same range. At $\times 1.05$ the initial issue therefore clips 2.34 times as many pixels as the current issue, and I_1 clips more than I_3 in 63.8% of the roots. What separates the images is the number of editor passes they have been through, not their generation label as such. Darkening produces no clipping, and none of the four endpoints loses status under it.

Table 21: Highlight statistics of the untransformed issues and unissued outputs over 480 roots (8-bit intensities; clipped-pixel shares in %). Brightening saturates every pixel above $255/c$.

			Clipped pixels (%) at brightness $\times c$			
Image	Mean	99th pct.	1.05	1.10	1.15	1.30
I_0 (initial issue)	109.06	241.53	3.83	5.79	7.98	14.62
I_1 (issue 1)	107.56	231.72	2.65	4.53	6.64	13.22
I_2 (issue 2)	108.99	227.60	1.99	3.92	5.95	12.58
I_3 (current issue)	106.94	221.84	1.64	3.12	4.86	11.26
U_3 (output after I_3)	108.50	221.92	1.68	3.36	5.24	12.03
U_p (branch)	110.38	227.65	1.89	3.55	5.77	13.14

Gamma reverses the order because it clips nothing: it is a monotone map within the 8-bit range, and what fails under it is the current-issue safeguard of Table 10, which withholds CURRENT when the current-issue evidence is insufficient. A SUPERSEDED decision passes through no such safeguard, so under gamma the initial issue at head 3 keeps 99.79% and 93.96% while the current issue keeps 75.00% and 66.67%, and the map is unchanged (93.13% at depth 3). Three limits apply. I_0–I_3 are different images, so the brightening gap is a property of the carrier images—a source set with fewer highlights would narrow it—rather than of the protocol. Clipping accounts for most but not all of the ordering: a model of status loss on the clipped share alone leaves a smaller but nonzero depth effect. And only the input highlight distribution was measured, not where the writer places its energy; a preference for high-contrast regions would add a second cause that the present measurement cannot separate.

E.5 BEYOND THE BOUNDARY, DESTRUCTION IS NOT AN ACCUSATION

Outside the envelope, a severely degraded issue and a complete-output edit produce the same observable pattern: issue evidence no longer decodes and the field is disrupted. The frozen router does not always separate the two, which is why heavy damage sends issued images to OPEN rather than to UNKNOWN. The map quantifies what survives such damage: at JPEG 30 the field still labels 85.00% of cells correctly and the newest slice keeps 90.25% F1. Whether that surviving history can also separate benign degradation from an unissued edit is untested here: an unissued output keeps its parent's regional history by construction, so a well-recovered map is compatible with either case. Reporting only payload survival would miss this collision, and reporting only OPEN recall would count it as a success; the paired matrix makes it explicit. Within the envelope, destruction never reads as an accusation: ordinary delivery preserves every status, unissued outputs remain OPEN, and the one failure direction is toward withholding.

F WHY SPATIAL CONTEXT RECOVERS CONSTANT-RATE HISTORY

Representing an expanding authorized history at a fixed local rate creates a predictable ambiguity: several legal generations can share the same local fingerprint. The spatial reader resolves that ambiguity by recalibrating candidate evidence and propagating reliable temporal anchors across coherent revision regions. We establish the full chain from the collision law, through independent-cell degradation, to spatial recovery.

F.1 Collision Probability and the Fixed-Slot Choice

For a b-bit keyed fingerprint and d other issued candidates, the ideal PRF model gives

$$p_{\mathrm{amb}}^{(b)}(d) = 1 - \left(1 - 2^{-b}\right)^{d}. \tag{13}$$

This per-region probability alone does not select a width: under a fixed carrier budget, wider fingerprints also produce fewer and coarser logical regions. Appendix L proves the resulting fixed-slot anchor law and evaluates the alternative rate–granularity allocations at $b \in \{1, 2, 3, 4, 8\}$. This appendix instead asks the downstream question: given the selected two-bit field, how does spatial inference recover the history compressed by its measured collisions?

F.2 Observed Ambiguity Follows the Collision Law

Table 22 compares the prediction with the frozen two-bit field on the three issued successor depths. Every row contains one own-head issue query over 480 independent roots (Figure 6). The empirical median differs from theory by at most 0.63 percentage points. The decline of cell-wise inference is therefore a direct representational effect: the hard local observation genuinely admits multiple legal generations.

Table 22: Collision formation and recovery by lineage depth. "Cells" is the median number of ambiguous cells among the 160 generation-field cells. Accuracies are medians over images and are restricted to ambiguous cells.

	Ambiguity			Ambiguous-cell accuracy ↑		
Depth	Theory (%)	Observed (%)	Cells	Local	Spatial	Gain ↑
1	25.00	24.38	39	46.67	88.64	+41.97
2	43.75	43.75	70	35.14	89.23	+54.09
3	57.81	58.13	93	25.56	88.27	+62.71

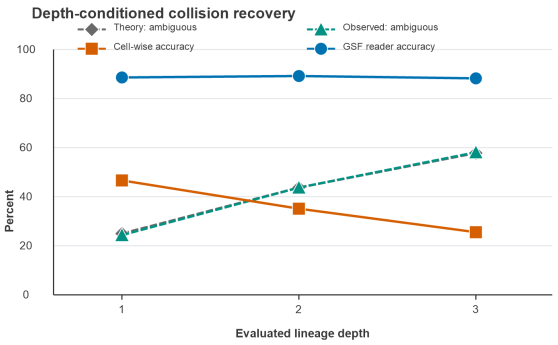

Figure 6: Depth-conditioned collision formation and recovery. Predicted and observed ambiguity increase together, independent cell-wise accuracy falls, and spatial recovery remains stable. Each point summarizes 480 issued successors.

As unique anchors become scarcer, independent-cell accuracy approaches chance among the colliding candidates. Spatial-reader accuracy remains near 89% through all three measured depths, so the absolute gain grows rather than shrinks. The candidate interface can score any legal fingerprint supplied to it; the concrete batched implementation pads candidates into eight planes and masks unused planes. The experiments use at most four legal generations. Because all candidate-plane weights are shared, the learned parameterization contains no generation-index-specific parameter. The evaluation uses four legal generations; Appendix L characterizes how the coding trade-off evolves at greater candidate depth.

F.3 READER INSTANTIATION

The frozen image decoder produces two bit logits and one confidence logit per cell. For legal candidate a, the exact summed local likelihood referenced in the main text is

$$u_{i,a} = \sum_{b=1}^{2} \log \sigma((2c_{L,a,i,b} - 1)\ell_{i,b}) \,. \tag{14}$$

The implementation uses five channels per legal candidate: mean two-bit log likelihood, the two individual bit log likelihoods, decoder confidence, and a logit-margin statistic. The batched tensor has shape $B \times 8 \times 5 \times 16 \times 16$, with a Boolean mask selecting legal candidates.

The deployed reader has 4,633 trainable parameters. A shared 1×1 stem maps five channels to 24 channels and applies GroupNorm and GELU. Three shared depthwise-separable residual blocks use 3×3 kernels with dilations 1, 2, and 4. Candidate-set context is formed only by masked mean and log-sum-exp pooling; a shared pointwise head returns one score per candidate and cell. There is no candidate embedding or candidate-order parameter.

These shared candidate operations and symmetric pooling give

$$\mathbf{p}_\psi(\pi Z) = \pi \mathbf{p}_\psi(Z), \tag{15}$$

Candidate permutation was tested by permuting the input candidate planes and undoing that permutation at the output. The maximum numerical difference was zero in the unit tests, including the pointwise and soft-spatial variants. Thus the implementation realizes Eq. (15), rather than learning that an early tensor index means an old generation.

Training data and selection. The deployed checkpoint was trained with a frozen writer and decoder on 160 distinct COCO val2017 sources: 64 for training, 32 for development, and 64 for sealed evaluation. The 2,048-example training cache contains 32 independently generated field/channel instances per training source. Generation fields use connected Voronoi, brush/polygon, rectangle/ellipse, holed, and disconnected same-label regions rather than independent cell labels. The fixed channel battery contains JPEG 80/95, resize factors 0.75/1.25, Gaussian blur, and light noise.

Training uses AdamW (Loshchilov & Hutter, 2019), learning rate 3×10^{-3}, betas $(0.9, 0.95)$, weight decay 5×10^{-3}, batch size 128, gradient clipping at one, and random candidate permutation. The loss is balanced cross-entropy plus Dice loss (Milletari et al., 2016) and a 0.5-weighted correct-key/wrong-key energy margin. Development metrics are evaluated every 100 steps; training stops after three non-improving checks. The frozen development criterion selects the step-200 checkpoint; the sealed split is not used for selection.

The reader-training sources are disjoint from the 480 evaluation roots.

F.4 DIFFERENT SPATIAL SOLVERS SHARE THE SAME UNDERLYING GAIN

Table 23 compares three ways of resolving the same candidate-likelihood field. The amortized soft reader is the deployed 4,633-parameter model above. A hard-partition reader learns affinities, constructs a region tree, and performs exact coded Bayesian decoding within each partition. A recurrent soft reader iteratively transports candidate evidence and additionally contains content and rejection heads. These are not three unrelated systems; they are alternative inference operators over the same keyed generation field.

Table 23: Alternative spatial inference operators on the same 480-root issued successors. Values are median ambiguous-cell accuracies. The first row is the independent-cell likelihood rule.

Inference operator	Depth-1 acc. ↑	Depth-2 acc. ↑	Depth-3 acc. ↑
Independent local likelihood	46.67	35.14	25.56
Hard partition + exact region decoding	84.21	82.02	80.23
Amortized soft spatial reader	88.64	**89.23**	88.27
Recurrent soft spatial reader	**88.89**	88.66	**89.01**

All spatial operators substantially outperform independent cells. The 4,633-parameter amortized reader matches the larger recurrent solver, while hard partitioning is consistently lower. This solver-family comparison triangulates the role of spatial evidence; the controlled study below isolates the two mechanisms that carry the gain.

F.5 Recalibration and Soft Coupling Carry the Gain

For the controlled mechanism study, all soft variants share a disjoint 80-source cache (56 train, 8 calibration, 16 development), a warm-started encoder, a region-balanced generation NLL, and 400 training steps. The 80 sources are disjoint from both the 480 evaluation roots and a separate 120-root reserve. Results are aggregated over the three own-head issued successor conditions (Table 24), so they must not be numerically merged with the depth-conditioned medians in Table 22.

Table 24: Controlled mechanism ladder on issued successors. "No spatial kernel" center-masks the learned depthwise kernels; shared GroupNorm and permutation-symmetric candidate context remain, so it isolates learned kernel coupling rather than every possible cross-cell statistic.

Mechanism	Ambiguous-cell accuracy (%) ↑	Increment ↑
Independent local likelihood	36.97	–
Learned evidence recalibration, no spatial kernel	67.48	+30.51
+ soft multiscale spatial kernels	87.83	+20.35
Recurrent soft coupling (three seeds)	87.70–88.11	$\leq +0.28$
Hard partition + exact decoding	80.78	-7.05

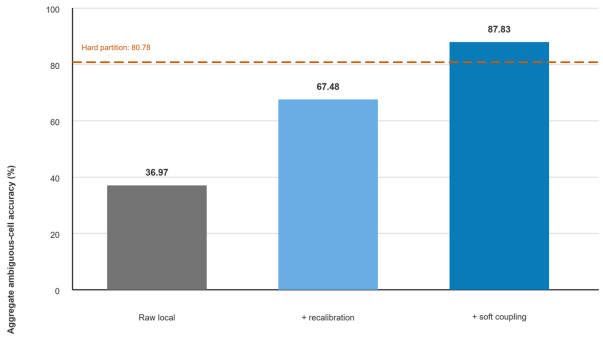

Figure 7: Reader-mechanism attribution on the 480-root transfer set. Shared pointwise recalibration and soft spatial coupling provide the two principal gains; the dashed line marks the hard-partition result.

The largest single improvement comes from learning how confidence, margin, and bitwise evidence should recalibrate the raw candidate likelihood (Figure 7). Spatial coupling supplies the second indispensable improvement by allowing reliable temporal anchors to support coherent neighboring cells. Recurrence changes the result by at most 0.28 points, showing that the compact feed-forward reader already captures the useful evidence flow. The independently trained solver family in Table 23 shows the same ordering, supporting soft evidence propagation over early hard partitioning.

F.6 Editing Changes the Reader's Task

Spatial inference is applied to issued fields whose carriers remain observable. Among unissued outputs whose generation carrier remains readable, the controlled reader recovers ambiguous cells with 92.11% accuracy. Where editing destroys that field, the joint verifier returns $\perp$ regionally and uses lineage and disruption evidence to identify the global OPEN/UNKNOWN transition instead of inventing a local history.

As an end-to-end consistency check, committed targets and decodable issued fingerprints agree on 99.988% of 133,373 cells. All images are retained, and the depth-conditioned conclusions are unchanged. The result reinforces the division of labor at the heart of GSF: the issuer's committed

field defines history, the image supplies evidence for it, and spatial inference reconstructs that history while editing-induced destruction opens the next lifecycle transition.

G ONE FIELD ANSWERS WHEN AND WHERE

The generation field is more than a device for improving ambiguous-cell accuracy. It changes what can be asked of a watermarked issue. Every label slice answers *when* the corresponding region was last authorized, while the newest slice answers *where* the current revision entered the lineage. Both answers come from the same recovered field and require neither a second localization watermark nor a separate localization head.

G.1 REPEATED REISSUANCE ACTUALLY PRODUCES MULTI-TIME IMAGES

The need for a generation map would be weak if each final image were dominated by a single generation. The opposite occurs. Over the full 16×16 logical revision grid, 463 of 480 final issues contain all four authorized labels, 16 contain three, and one contains two. Table 25 shows how much of the final field remains attributable to each time. The latest generation is largest on average, but every earlier generation remains present in almost all final fields.

Table 25: Number of logical cells carrying each last-authorized generation in the final ground-truth field. Quartiles and extrema are over 480 roots on the 256-cell logical grid; "Present" counts fields with at least one such cell.

Label	Present	Minimum	Q1	Median	Q3	Maximum
g_0	480	1	18	33	52	168
g_1	466	0	9	19	33	102
g_2	476	0	25	44	71	153
g_3	480	64	126	152	170	220

This coexistence follows directly from inheritance. A reissue assigns the new label only on the issuer-committed revision scope and freshly re-embeds the inherited labels elsewhere. The committed scopes are substantial but do not overwrite the whole field: their median sizes on the 256-cell logical grid are 147, 147, and 152 cells at depths 1–3, with respective interquartile ranges 120.5–168, 119–168, and 126–170. Repeated revision therefore creates a spatial mosaic of authorized times rather than a sequence of whole-image version numbers.

The logical field is defined on all 256 cells. The concrete carrier jointly allocates 96 cells to lineage/issue evidence and 160 cells to generation fingerprints. Recovery metrics evaluate all 160 history-bearing cells. Generation-specific IoU is reported whenever the corresponding label is represented in that evaluated field.

G.2 SPATIAL INFERENCE RECOVERS THE TEMPORAL PARTITION

Table 26 moves from ambiguous cells to the scientific object of interest: the complete regional-history map. Independent local likelihood deteriorates sharply as more historical labels coexist. Spatial inference maintains 93.13% full-field accuracy and 78.96% mIoU even at depth three, where a median 58.13% of field cells are locally ambiguous. Figure 8 shows the transition from independent to spatial readout on a representative four-generation successor; Figures 9 and 10 follow two further lineages issue by issue, with the unissued output that each reissue approved and the final unissued output.

Table 26: From collision recovery to generation-map recovery. Map metrics use all 160 generation cells. Scope metrics use the newest-generation slice. Entries are medians over 480 issued successors at each depth.

	Independent local likelihood				Spatial generation-field reader			
Depth	Map acc. ↑	Map mIoU ↑	Newest IoU ↑	Newest F1 ↑	Map acc. ↑	Map mIoU ↑	Newest IoU ↑	Newest F1 ↑
1	86.88	75.82	75.00	85.71	97.19	93.87	94.92	97.39
2	71.88	52.61	56.09	71.87	95.00	87.27	94.09	96.95
3	57.50	36.76	42.19	59.34	93.13	78.96	92.86	96.30

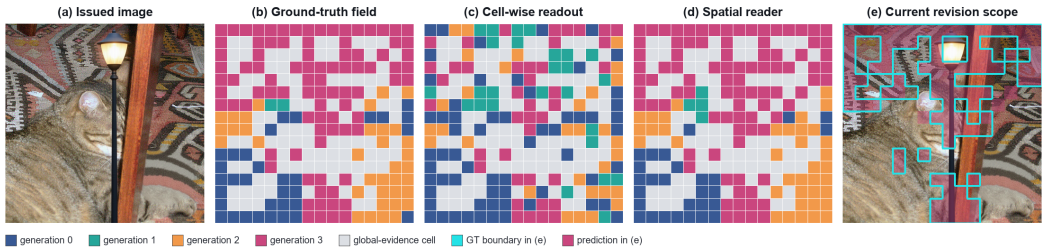

Figure 8: Spatial inference recovers a representative four-generation field. Independent cell-wise readout confuses historical labels, whereas the spatial reader restores the field used for both temporal attribution and localization. In (e), magenta is $\widehat{A}_3 = \{i : \widehat{\tau}_3(i) = 3\}$ and cyan traces the committed ground-truth A_3. Gray cells carry global lineage and issue evidence and are excluded from regional metrics.

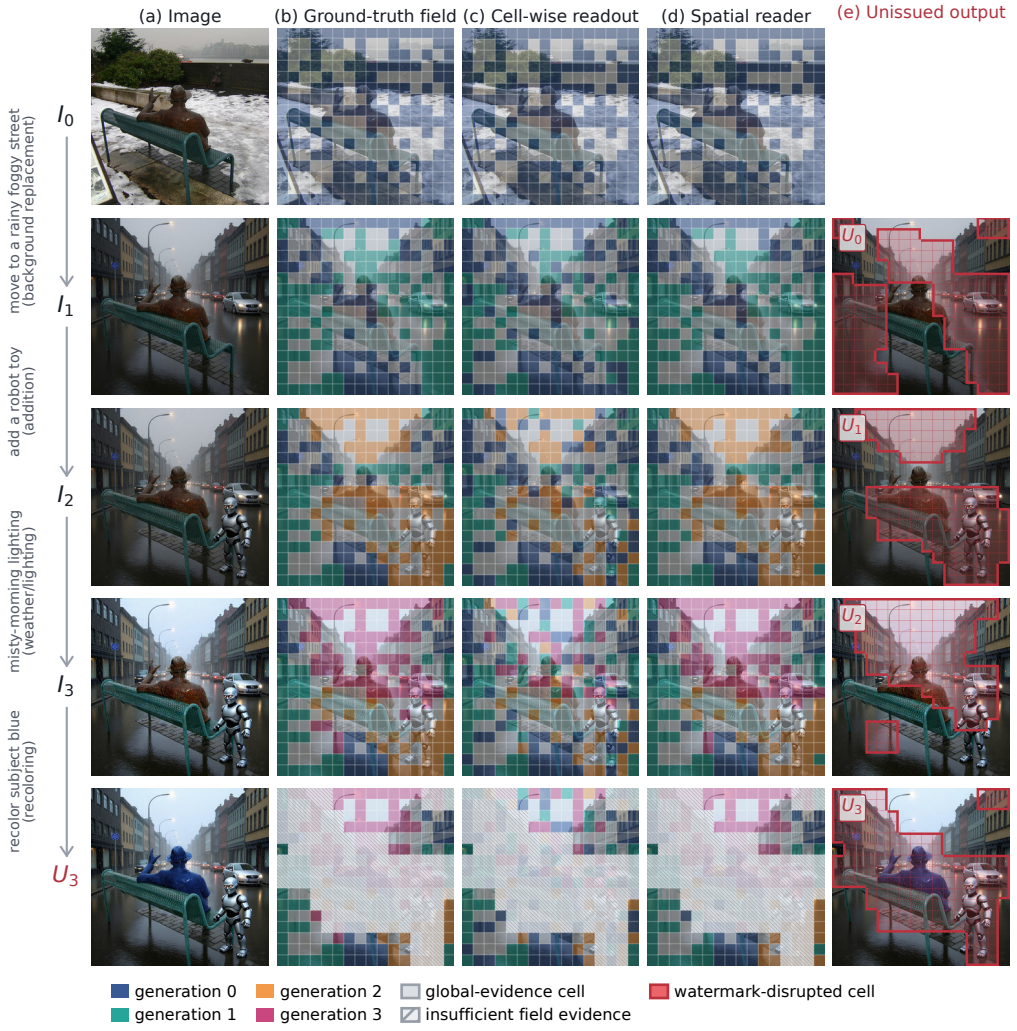

Figure 9: A second lineage followed issue by issue. Rows are the issues I_0–I_3 and the final unissued output U_3; the label between two rows is the edit that produced the lower one. Columns: (a) the image; (b) the ground-truth field, for U_3 the field of its parent I_3; (c) the cell-wise readout; (d) the spatial reader; (e) the unissued output that the reissue in that row approved, with its watermark-disrupted cells outlined, and in the last row U_3 itself. Disrupted cells of U_3 carry insufficient field evidence and are hatched in (b)–(d). Reissues by FLUX.1 Kontext throughout; U_3 by Qwen Image Edit.

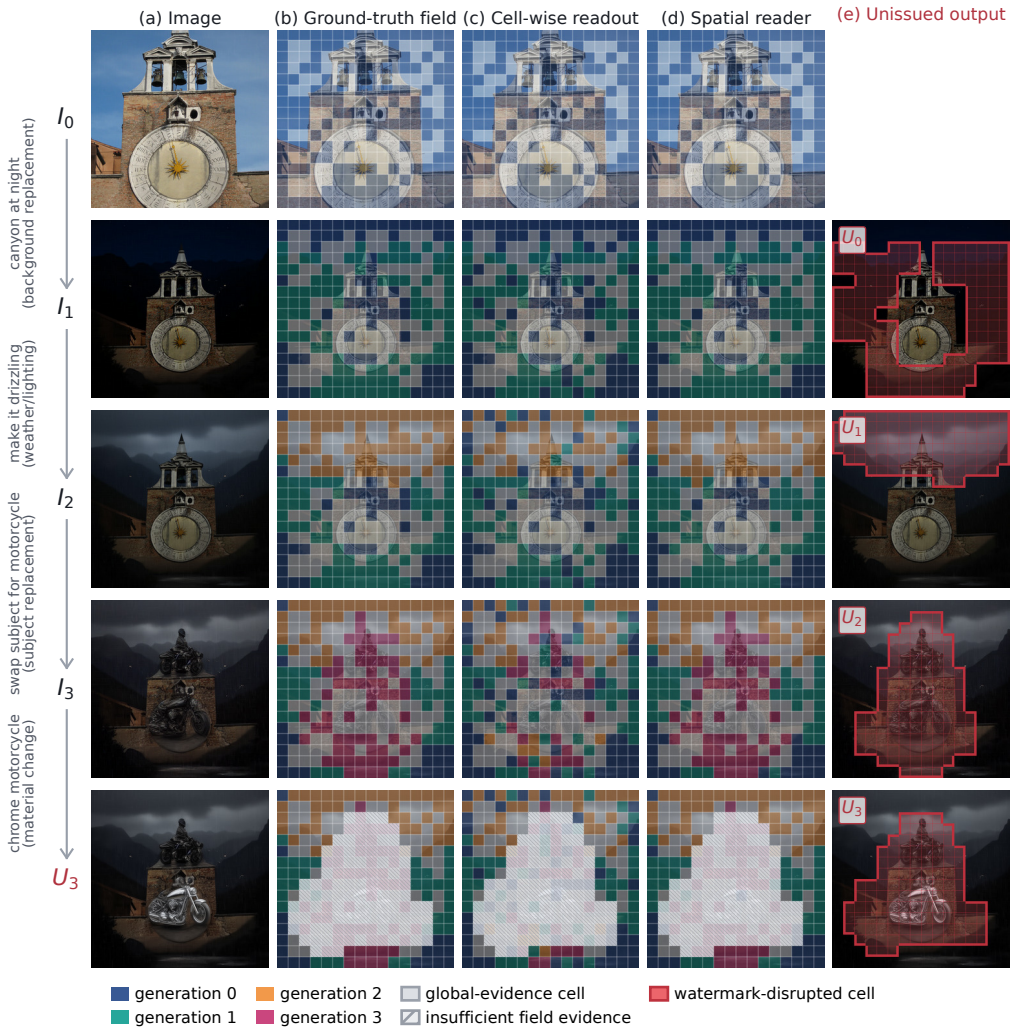

Figure 10: A third lineage followed issue by issue. Rows are the issues I_0–I_3 and the final unissued output U_3; the label between two rows is the edit that produced the lower one. Columns: (a) the image; (b) the ground-truth field, for U_3 the field of its parent I_3; (c) the cell-wise readout; (d) the spatial reader; (e) the unissued output that the reissue in that row approved, with its watermark-disrupted cells outlined, and in the last row U_3 itself. Disrupted cells of U_3 carry insufficient field evidence and are hatched in (b)–(d). Reissues by FLUX.1 Kontext, Qwen Image Edit, and InstructPix2Pix in turn; U_3 by Qwen Image Edit.

The aggregate decline is asymmetric rather than uniform across the field. From depth one to depth three, generation-field mIoU falls from 93.87% to 78.96%, whereas newest-slice IoU moves only from 94.92% to 92.86%—an end-to-end change of 2.06 percentage points. Class-balanced overlap exposes the harder part of the problem: older labels occupy smaller surviving regions, so boundary errors weigh more heavily as those regions fragment and shrink. Table 27 localizes this loss. At depth three, the current g_3 slice retains 92.86% IoU, while the historical g_0, g_1, and g_2 slices reach 81.48%, 72.57%, and 81.32%, respectively. Within the tested depths, the loss is therefore concentrated in the completeness of fine-grained historical reconstruction rather than a comparable destabilization of the current committed scope.

G.3 The Newest Slice Is the Current Authorized Scope

For a committed successor g, the update law gives

$$A_g = \{i : \tau_g(i) = g\}, \qquad \widehat{A}_g = \{i : \widehat{\tau}_g(i) = g\}. \tag{16}$$

Table 27: Presence-conditioned IoU of every authorized-generation slice under spatial inference. A dash denotes a generation that does not yet exist at that depth. Entries are median percentages.

	Presence-conditioned IoU (%) ↑			
Depth	g_0	g_1	g_2	g_3
1	93.99	94.92	–	–
2	89.32	83.33	94.09	–
3	81.48	72.57	81.32	92.86

Consequently, reading all slices yields the regional authorization history, whereas selecting only the newest slice yields the current committed revision scope. Table 26 shows that its IoU remains 94.92%, 94.09%, and 92.86% at depths 1–3, while its F1 remains above 96% throughout. This stability is the spatial half of the joint task: the map does not merely say that a revision happened; it preserves the present issuer-committed scope even as older authorization fragments accumulate, and attributes every observable region to the authorization event that most recently changed it.

In the evaluation, the proposed scope is constructed from blockwise RGB difference between the issued parent and the editor's complete return, using Eq. (12); the issuer then commits that proposal. Thus the newest slice localizes the *committed, pixel-difference-derived revision support*: the blockwise region through which the revision entered authorized history. This authorization target is distinct from an editor's latent pixelwise semantic mask. In another deployment, spatial watermark disruption or authenticated edit records may propose the scope; issuer commitment gives any such proposal its authority.

Ordinary edit localization asks which pixels differ in one snapshot. The generation map instead records which authorized event currently governs each region. Its newest slice provides the current revision support, while its older slices retain the temporal attribution that a binary edit mask discards.

G.4 Status and Regional History Are Recovered Together

Table 28 pairs status and regional history on the same issued successors and the same live head, showing that one verifier supplies both answers from the same image.

Table 28: Paired global and regional results on the same own-head successor queries. Every depth is one query condition over 480 roots. Regional entries are the medians defined above; global accuracy is the fraction predicted CURRENT.

Depth	Predicted CURRENT (%) ↑	Ambiguous-cell acc. (%) ↑	Map accuracy (%) ↑	Newest IoU (%) ↑
1	99.79	88.64	97.19	94.92
2	99.79	89.23	95.00	94.09
3	98.96	88.27	93.13	92.86

Across these 3×480 successors, 1,433 are globally recognized as the live issue. On the same queries, the regional field retains high full-map and newest-scope accuracy despite increasing local collisions. The resulting object is therefore not a version number plus an unrelated mask. It is a single spatiotemporal authorization field: one field answers when and where, and continuity is maintained jointly at the issue level and the regional-history level.

H Canonicalized Verification Preserves Multi-Resolution Delivery

The writer is trained and the keyed 16×16 field is defined on a canonical 512×512 canvas. Images may nevertheless circulate at different delivery resolutions: verification restores the canonical canvas before reading the field. Across intermediate sizes from 768 to 2048 pixels, this path preserves every tested issue and can even make fixed-pixel delivery operations milder relative to the carrier.

H.1 Resolution Round Trips Preserve the Carrier

We write at 512, resize to an intermediate square canvas, and restore 512 before verification. The supporting set contains 60 issued images. Twelve round-trip settings cover intermediate sizes 768,

1024, 1536, and 2048 and PIL bilinear, bicubic, and Lanczos kernels, together with area downsampling (Table 29). No network is retrained and the verifier is unchanged.

Table 29: The physical carrier survives ordinary resize round trips. The carrier score is the median absolute two-bit decoder margin aggregated over the field; the ratio is paired to the clean score. Every tested resampling pair in a row obtains the stated result on all 60 issues.

Intermediate canvas	Kernel pairs	Carrier score ↑	Clean ratio ↑	CURRENT ↑
None (clean)	1	17.999	1.000	60/60
768	2	18.013–18.042	1.002–1.003	60/60
1024	6	17.929–18.040	0.996–1.002	60/60
1536	2	18.016–18.030	1.001	60/60
2048	2	18.030–18.048	1.001–1.002	60/60

The median number of disrupted cells is zero in every setting. This result rules out the intuitive but incorrect explanation that the carrier lives near the pixel Nyquist limit. Its semantic lattice is 16×16, so its characteristic scale is approximately 32 pixels on the canonical image. The tested linear resamplers preserve structure at that scale even when they alter individual pixels. This also explains why benign resampling and AI regeneration must not be conflated: the former filters an existing block-scale signal, whereas the latter can replace the image-wide carrier evidence while producing a new complete output.

H.2 CANONICAL COORDINATES PRESERVE KEYED CELL IDENTITY

Canonicalization is part of the spatial interface rather than an image-quality repair. It restores the same 16×16 lattice on which every keyed cell index was written, so resizing changes the delivery canvas without changing the semantic identity of a region. The zero disrupted-cell median in Table 29 confirms that ordinary resampling preserves the underlying block-scale evidence.

H.3 A CANONICAL HUB CAN IMPROVE DELIVERY ROBUSTNESS

We next compare a native-512 delivery path with a path that distributes the same issued image at 1024 and canonicalizes it to 512 for verification. Both arms use the same 60 roots and the same six benign operations. Table 30 shows that the added round trip introduces no clean penalty and improves the weakest fixed-pixel transformations in this set.

Table 30: Matched benign verification at native 512 and through a 1024-distribution/512-verification hub. Entries are CURRENT rates over 60 issues per condition.

Condition	Native 512 ↑	Canonical hub ↑	Difference ↑
Clean	1.0000	1.0000	0.0000
JPEG 75	1.0000	1.0000	0.0000
Downsample/upsample	1.0000	1.0000	0.0000
Crop 90% and resize	0.8000	0.9333	+0.1333
Rotate 2°	1.0000	1.0000	0.0000
Screenshot simulation	0.9667	1.0000	+0.0333
Pooled	0.9611	0.9889	+0.0278

The crop improvement is not explained by a different recovered scale: the median alignment scale is 1.1087 in both arms. A consistent explanation is relative operator scale. Distribution at 1024 doubles the carrier block to approximately 64 pixels, while JPEG blocks and interpolation kernels remain fixed in pixel units; those operations are therefore milder relative to the carrier. The pure geometric rotation, for which this argument predicts little change, is unchanged.

H.4 ALIGNMENT LIVES IN THE SAME CANONICAL COORDINATES

The blind alignment search in Appendix B handles isotropic scale, rotation, and local translation before field decoding. Both alignment and resolution normalization therefore operate in the same canonical coordinates. Across all tested delivery resolutions, the physical evidence survives resampling and canonicalization restores the semantic identity of every keyed spatial cell.

I COMPLETE OUTPUTS CHANGE THE LOCALIZATION ENDPOINT

The complete-output workflow changes which image is retained rather than the size of the inpainting mask or the strength of a post-processing attack. A paste-back endpoint preserves the watermarked container pixel for pixel outside the edit mask; a retained-output endpoint keeps the editor's complete RGB return, including its reconstruction of the nominally unmodified background. This appendix isolates that single endpoint choice and shows why snapshot localization and status verification answer different questions under it.

The controlled comparison first verifies each snapshot localizer in its native paste-back protocol, then changes only the retained endpoint. It thereby isolates the physical premise removed by complete-output editing.

I.1 MATCHED ENDPOINT CONSTRUCTION

The matched control uses all 1,000 images and masks in valAGE (Zhang et al., 2024). The median mask area is 0.0918. For each carrier, source, mask, prompt, and seed, one inpainting call produces an edited image Y. We then form

$$X_{\text{paste}} = (1 - M) \odot X_{\text{wm}} + M \odot Y, \tag{17}$$

$$X_{\text{full}} = Y, \tag{18}$$

where X_{wm} is that method's watermarked container and M is the provided mask. Thus the two endpoints within each method have identical generated content and differ only in whether the protected background is restored. Active methods require their own containers, so pixels are not shared across methods; the image set, masks, editor, prompts, and seed schedule are shared.

All arms use the same SD-1.5 inpainting editor (Rombach et al., 2022) with DDIM (Song et al., 2021), 50 steps, guidance scale 7.5, and the safety checker disabled. Sharing the editor makes retained endpoint the controlled variable; the native-endpoint results first anchor both released method implementations (Table 31).

Table 31: External localization configurations in the matched endpoint control. EditGuard uses its released mask rule; OmniGuard uses the standard 0.5 decision point for its sigmoid map.

Method	Released weight	Localization signal	Native threshold
EditGuard	clean	revealed blue residual	0.2
EditGuard	degradation-trained	revealed blue residual	0.4
OmniGuard	DTE checkpoint 1500	sigmoid tamper probability	0.5

EditGuard is loaded strictly with its released clean and degradation-trained weights; both recover the embedded 64-bit message with zero BER on unedited containers. Its published mask rule, thresholding whether any revealed channel differs from the blue reference, is used without learned post-processing. OmniGuard (Zhang et al., 2025) is loaded with complete checkpoint coverage and uses its released fixed secret image, padding/normalization pipeline, and DTE output. Its clean-container artifact reconstruction and negligible false positive area provide an independent forward/reverse-path check.

I.2 NATIVE SUCCESS ISOLATES THE ENDPOINT EFFECT

Table 32 first evaluates the paste-back image, for which the supplied mask remains the exact composition support. Both external localizers reproduce strong native behavior: EditGuard reaches 0.9718/0.9519 F1 with its two released weights, and OmniGuard reaches 0.9847. Their average precisions are 0.9968, 0.9949, and 0.9994, respectively. This strong native performance anchors the checkpoints, mask convention, and operating rules before the retained-output endpoint is evaluated.

Once the full output is retained, the two localizers move in opposite directions. EditGuard predicts the entire image: its F1 equals the all-positive identity $2|M|/(1 + |M|)$ rather than adding localization information. OmniGuard predicts almost nothing. The pattern holds across a wide fixed-threshold sweep: on the retained outputs EditGuard's revealed residual stays above threshold over the whole image while OmniGuard's soft map has negligible support, and the carrier diagnostics in Table 33 give the physical cause.

Table 32: Changing only the retained endpoint changes the output of snapshot localizers. F1 and predicted area are medians over 1,000 images. The mask prior is 0.0918, so an all-positive prediction has F1 0.1681.

	Native paste-back		Retained complete output	
Method	F1 ↑	Pred. area	F1 ↑	Pred. area
EditGuard, clean weight	0.9718	0.0962	0.1681	1.0000
EditGuard, degradation weight	0.9519	0.0850	0.1681	1.0000
OmniGuard	0.9847	0.0935	0.0000	0.0027

Table 33: Carrier evidence under clean, paste-back, and retained-complete endpoints. BER is better when lower; artifact PSNR is relative to OmniGuard's secret localization image; GSF disruption is the fraction of unreadable carrier cells.

Diagnostic	Clean	Paste-back	Complete output
EditGuard clean-weight message BER ↓	0.0000	0.0000	0.5000
EditGuard degradation-weight message BER ↓	0.0000	0.0000	0.5156
OmniGuard artifact PSNR (dB) ↑	38.66	17.35	6.67
GSF disrupted-cell fraction	0.0000	0.0837	0.8778

Paste-back leaves enough predecessor carrier outside M for the native localizers to compare intact and disrupted evidence. Complete-output generation instead randomizes EditGuard's decoded message and collapses OmniGuard's recovered secret image from 38.66 to 6.67 dB. The localization outputs therefore reflect predecessor-evidence loss, not evidence that the retained image is unedited.

I.3 The Original Mask No Longer Describes Every Changed Pixel

For X_{paste}, pixels outside M are copied from the container by construction, so M is a valid pixel-composition target. For X_{full}, the editor has returned new RGB values outside M as well. The original mask still identifies the requested inpainting support, but it no longer describes every changed output pixel. Consequently, the retained-output F1 values above diagnose how a snapshot localizer behaves when its native background-retention premise is removed; they are not a semantic ranking of complete-output localization methods.

GSF addresses the reissuance that follows rather than reusing the inpainting mask as its target. The event graph identifies the complete return as an unissued transition; after review, the issuer commits an observed revision scope and writes a fresh successor carrying the inherited history.

I.4 The Effect Persists Across Complete-Output Editors

A supporting control replaces the inpainting editor with the four complete-output editor families used in the reissuance evaluation. It covers 12 roots (two FLUX, five InstructPix2Pix, three Qwen Image Edit, and two SDXL-Turbo instances) under matched prompts and seeds, reproducing the endpoint mechanism in every editor family (Table 34).

Table 34: Supporting complete-output control over 12 roots and four editor families. The trivial F1 differs slightly across carriers because each method writes its own container and therefore induces its own output.

Method	Carrier diagnostic	Pred. area	F1 ↑	Output form
EditGuard, clean weight	BER ↓: 0.5391	1.0000	0.7495	all positive
EditGuard, degradation weight	BER ↓: 0.4844	1.0000	0.7532	all positive
OmniGuard	artifact PSNR ↑: 6.57 dB	0.0039	0.0096	nearly empty

The matched 1,000-image experiment supplies the causal endpoint comparison; the four-editor control shows that its physical explanation is not peculiar to a single inpainting implementation. The consequence is that watermark destruction has acquired a meaning in the protocol: it marks an unresolved, unissued transition, and authorized reissuance converts that transition into a new readable issue carrying forward the regional history.

J Fidelity, Statistical Design, and System Cost

This appendix quantifies image fidelity, identifies the independent experimental unit, and separates constant image-resident rate from registry growth. Together, these measurements show that status and regional-history results are obtained with a high-fidelity, compact carrier and root-level statistical accounting.

J.1 Initial Reissuable Issues Remain High Fidelity

For every one of the 480 evaluation roots, the initial issue is written from its canonical source and compared with that source in 8-bit RGB. Table 35 reports the complete PSNR distribution. The median is 46.694 dB and the tenth percentile remains 42.504 dB. Thus the results do not rely on a visibly aggressive initial carrier.

Table 35: Initial-issue fidelity over 480 roots. PSNR is computed between the canonical source and its newly written issue. Quantiles use linear interpolation between order statistics.

	PSNR (dB) ↑					
n	Minimum	P10	Mean	Median	P90	Maximum
480	36.850	42.504	45.973	**46.694**	48.141	48.722

The interquartile range is 45.082–47.598 dB. A 10,000-replicate root bootstrap (seed 20260824) gives a 95% interval of $[46.498, 46.816]$ dB for the median and $[45.766, 46.175]$ dB for the mean. The separate 24-image physical-channel validation set in Table 9 additionally records median LPIPS 0.003431 at 45.646 dB PSNR, confirming that the learned carrier remains visually light while supporting the complete reissuance protocol.

J.2 The Independent Experimental Unit Is the Root

The generation lattice yields many cells per image, and the same root is queried at several live heads. Neither repetition creates new independent images. Table 36 records the unit and aggregation behind each headline result. Exact count tables are reported when available; when an uncertainty interval is used, resampling is paired or clustered at the root, never at the cell.

Table 36: Statistical interpretation of the main evidence. Query conditions and post-processing settings are repeated measurements within a root, not additional independent image draws.

Result	Unit	Denominator	Aggregation and dependence
Status of issued images	root–head query	4×480 own-head and 6×480 later-head queries	Pooled fixed-condition accuracy; roots cluster repeated heads
Status of unissued outputs	root–head query	3×480 pre-reissue, 1×480 final-unissued, and four earlier-issue conditions ($249 + 249 + 231 + 480$)	Pooled recall; event truth fixed before evidence; roots cluster repeated queries
Post-processing robustness	root–condition query	The clean images and 31 prespecified transformations on reused roots	Per-condition rates followed by median/minimum stress summaries; conditions are not iid draws
Collision recovery	successor root	480 successors independently at each tested depth	Ambiguous-cell accuracy is computed within image, then summarized across roots
Generation maps	successor root	480 successors at each tested depth	Per-image accuracy/IoU, then median; absent labels do not create artificial zero-IoU classes
Matched endpoint control	image	1,000 valAGE images	Paired by source, mask, editor configuration, and seed within each carrier

This accounting is especially important for the reader result. Reporting a confidence interval over tens of thousands of spatial cells would make the interval spuriously narrow because neighboring cells share content, edit support, lineage, and decoder noise. The paper instead reports per-successor medians at each depth and keeps the 480-root aggregate mechanism ablation separate from those depth-conditioned values. Similarly, the joint global and regional result uses the same successor/head pairs but does not condition the regional metric on a correct global prediction.

J.3 Fresh Authority, Offline Regional Readout, and Registry Cost

Two facts must be fetched live: whether the lineage is still active and which issue is current. Appendix A establishes why this is structural. A fixed bitmap and entirely fixed verifier state cannot express a later change from CURRENT to SUPERSEDED; the transition requires a trusted mutable coordinate, and the live head supplies it. The registry's verification interface is exactly that head and the lineage's label set $\mathcal{G}_L$ (Section 3). The label set is append-only and changes only at issuance, so a verifier can cache it; the registry is needed live only for h.

Regional history is read from the copy. The spatial reader needs the image, the key K, the lineage identifier decoded from the image, and the label set that defines its candidate planes (Section 4.3); it does not consult the head, and it does not consult the committed scopes A_g. Those scopes remain in the registry as the authorization records that a recovered label points to, but the readout does not depend on their being retained, available, or shared, so a partner institution or a forensic unit entrusted with K reads each region's last authorized revision from the pixels alone. For an unissued output there is no registry record at all: its inherited history and the disruption that marks the edit exist only in the copy, which is the evidence Sections 5.2 and 5.4 read.

The costs follow this placement (Table 37). GSF keeps every issued image at a fixed 512-slot payload, independent of lineage depth. Once a copy has circulated, later issuance never requires locating, enlarging, or rewriting it. Each new issue instead appends one authenticated issue record and advances the live head, which places growing authorization history in the component designed to authenticate, audit, and update it: the registry.

Table 37: Placement of model, payload, authorization-state, and query costs in the evaluated instantiation. n denotes the number of issues in a lineage.

Component	Evaluated cost	Scaling or role
Fixed distributed representation		
Writer	596,355 parameters	one full rewrite per issuance
Decoder	2,329,155 parameters	one canonical evidence field
Spatial reader	4,633 parameters	shared over candidate and spatial axes
Learned total	2,930,143 parameters	independent of lineage depth
Image-resident payload	512 raw binary slots	2 slots per cell; depth independent
Authenticated authorization state		
Mutable live head	one authenticated issue label or inactive marker	constant record count; value uses $O(\log n)$ bits
Issue registry	one authenticated record per issue	$O(n)$ records preserving the complete issue-level audit history
Current verification work		
Candidate evaluation	n candidate checks	query-time cost of the exhaustive scan in the evaluated verifier

The evaluation uses four legal generations, while shared candidate weights provide a variable-set interface. Each appended issue record extends the authenticated audit trail without changing any circulated image payload or learned model. GSF therefore concentrates lineage growth in an updateable registry rather than distributing it across immutable image copies: the representation carried by every copy remains fixed, while authorization history remains centrally auditable.

The evaluated verifier exhaustively scores the candidate set $\mathcal{G}_L$, giving $O(n)$ query work. Table 37 reports this linear scan as an implementation property, separately from the $O(n)$ persistent audit history. Candidate retrieval remains an implementation dimension rather than a consequence of the constant-rate field.

J.4 Evidence Map

Table 38 connects each experimental component to the claim it establishes.

The interface also extends beyond AI editing: conventional layer- or pixel-based operations can propose a revision scope and enter the same issuer-commitment workflow. Complete-output AI editors provide the demanding instantiation evaluated here.

Table 38: How the experimental evidence supports the paper's claims.

Axis	Evidence	Established result
Status semantics	Current/superseded/open behavior across repeated complete-output editing, branching, and head changes on 480 roots	A live head expires old issues, while reissuance separates unissued edits from issued successors
Regional history	Four coexisting labels, depth-conditioned collisions, and generation-map recovery	One constant-rate field recovers temporal attribution and current revision support
Resolution	Canonicalized delivery from 768 through 2048 pixels	Resolution changes preserve keyed cell identity and all tested issues
Post-processing	Paired issued/unissued queries under 31 frozen transformations	Routine delivery preserves both status decisions; heavier stress reveals their distinct evidence curves
Authorized scope	Issuer commitment of a block-wise pixel-difference proposal derived from each complete editor return	The newest generation slice localizes where the current revision entered authorized history
External controls	Matched native-paste-back versus retained-output endpoint under a shared SD-1.5 inpainting setup	Complete outputs remove the predecessor-residual premise used by snapshot localizers

J.5 Artifact Traceability and Evidence Chain

The anonymous manifest records each artifact's digest, version, producing script, and table/figure dependencies. Together with the frozen component inventory in Appendix A, it traces every aggregate back to its root, issuance event, and verifier output.

Taken together, the evidence forms one chain rather than a collection of independent benchmarks: complete-output editing opens an unresolved transition; authenticated reissuance restores a readable issue without requiring pixelwise survival of its predecessor; and the same constant-rate field recovers where each region last entered the authorized history. Continuity is thus maintained jointly at the status and regional-history levels.

K Two Same-Rate Codes for the Field: Direct Labels or Keyed Compatibility

The generation field of Section 4.2 can be written at the same two-bit rate with either of two codes, both implemented and evaluated here on the same 480 roots. A direct code gives each generation label its own two-bit symbol: for a fixed alphabet of at most four labels the assignment is a bijection, and an intact symbol maps back to its label without any candidate collision. The keyed field instead truncates a keyed fingerprint of (K, L, a, i) to two bits, which deliberately permits position-dependent local collisions among candidates.

The two codes answer different requirements. Where the label alphabet is fixed in advance and holds at most four generations, the direct code is the better local code, and Appendix K.3 measures by how much. GSF is designed for a candidate set that grows after the carrier is frozen, for a code-level signal that a readable symbol matches no issued candidate, and for spatial anchors that survive local ambiguity. A direct code spends the entire two-bit alphabet on the closed lookup: its fifth label is not occasionally ambiguous but indistinguishable from an earlier label everywhere, and a damaged symbol is still read as one of the four labels. The choice is therefore not between collisions and no collisions but between deterministic global aliasing and stochastic local ambiguity. The former erases a distinction from the field; the latter is detectable and can leave distinguishing evidence at other positions.

K.1 One Rate, Two Different Representational Commitments

Let a denote an issued generation, i a spatial cell, and $q_{L,i} = \mathrm{Trunc}_2 F_K(L\|i)$. We compare

$$c^{\mathrm{key}}_{a,i} = \mathrm{Trunc}_2 F_K(L\|a\|i), \tag{19}$$

$$c^{\mathrm{dir}}_{a,i} = \mathrm{bin}_2(a \bmod 4), \tag{20}$$

$$c^{\mathrm{pdir}}_{a,i} = c^{\mathrm{dir}}_{a,i} \oplus q_{L,i}. \tag{21}$$

The first is GSF's candidate-compatibility fingerprint. The second is a position-invariant two-bit version tag; on the alphabet $\{0, 1, 2, 3\}$ used throughout, $a \bmod 4$ is a bijection, and any label outside that alphabet receives the symbol of a label inside it at every cell. The third is a diagnostic keyed-permuted direct control: an XOR pad hides the visible label-to-symbol mapping while preserving the direct code's generation aliases. In particular, adding such a pad does not distinguish a from $a + 4$.

All three assignments use the same 320 physical binary slots and the same 160 two-bit logical cells. Given a hard symbol z_i, exact candidate matching returns a label only when $\{a : c_{a,i} = z_i\}$ is a singleton. A soft local rule instead maximizes the candidate likelihood in Eq. (14); a coverage-matched selective rule abstains at a threshold chosen to equal the keyed exact rule's issued-cell answer rate. The GSF spatial reader receives the same candidate-likelihood field but may use evidence outside i. Abstentions count as errors in every unconditional accuracy reported below.

Table 39 summarizes why equal code rate does not imply equal field semantics. The direct code is optimized for the first row; the keyed field is designed around the remaining rows as well.

Table 39: Representational commitments at the same two-bit rate. Candidate growth means adding issued generations while retaining the existing carrier interface. The rows state properties of the codes themselves; a readout rule may add a confidence threshold to either code (Appendix K.3).

Property	Direct two-bit tag	GSF keyed field
Intact labels 0–3	Collision-free bijection	Predictable local stochastic ambiguity
Candidate growth	a and $a + 4$ alias at every cell	Same interface; unique anchors decrease with candidate count
Encoded relation	Generation modulo four	Compatibility for (K, L, a, i)
Damaged symbol	Every two-bit value remains a legal label	Exact matching can return no match or multiple matches
Spatial evidence	A global alias leaves no distinguishing position	A pair may collide here and differ elsewhere
Wider code	Postpones alias but reduces fixed-slot spatial resolution	Retains the two-bit, 160-cell operating point

K.2 A Code-Only Counterfactual on Fixed Physical Evidence

To isolate this representational difference, the code-only comparison holds the decoded physical evidence fixed. If slot j was written with sign $s_j \in \{-1, +1\}$ and has decoder logit ℓ_j, its truth-aligned evidence is $e_j = s_j \ell_j$. For an alternative assignment with sign s'_j, the counterfactual logit is

$$\ell'_j = s'_j e_j. \tag{22}$$

Equation (22) preserves evidence magnitude, revision field, slot location, and image-level query while changing only the code assignment. The construction favors the direct code in one respect: it does not charge that code for the carrier's co-adaptation to the keyed code and its alignment objective, which the physical-recoding control of Appendix K.7 measures separately. Reinstating the native keyed signs reproduces the original measurements exactly.

The counterfactual covers all 17 issued, unissued, and branch query conditions over the same 480 roots. Complete per-condition results are provided in the machine-readable artifacts; here we retain the comparisons needed to separate a closed-set lookup from a reissuable field.

K.3 Direct Labels Are the Better Local Code on a Fixed Four-Label Alphabet

On issued images, the keyed exact rule answers 57.8876% of cells and obtains 57.8806% unconditional accuracy; 42.1124% of cells admit more than one candidate. The direct assignment is injective

over the four labels in this experiment and reaches 0.99983 unconditional issued-cell accuracy. Its soft local rule reaches 0.99990. When the direct rule is thresholded to match the keyed exact rule's answer rate, its selective error is 4.2×10^{-5}, compared with 1.33×10^{-4} for keyed exact matching; a confidence threshold thus makes the direct code selective as well.

The keyed ambiguity is not unexplained decoder noise. With one, two, and three other candidates, its cell-pooled ambiguous fractions are 0.248698, 0.436289, and 0.578138, closely matching $1 - (3/4)^d = 0.25, 0.4375, 0.578125$; the largest deviation is 0.0013. The 24.38%, 43.75%, and 58.13% values in Table 2 are per-image medians and therefore use a different aggregation.

The direct code is thus the better local code for this closed, intact, four-label question, as a bijection over a fully known alphabet should be. The comparison rests on two assumptions that the following sections release in turn: that the alphabet is fixed in advance, and that the observed symbol remains meaningful.

K.4 Candidate Growth and Local Rejection

The depth failure is global, not local. For direct b-bit coding, $c_a^{\mathrm{dir}} = c_{a'}^{\mathrm{dir}}$ whenever $a \equiv a' \pmod{2^b}$. At two bits,

$$c_{a,i}^{\mathrm{dir}} = c_{a+4,i}^{\mathrm{dir}} \qquad \text{for every cell } i. \tag{23}$$

Labels 0 and 4 therefore do not merely collide in an unlucky subset of cells; the field contains no position at which they differ. No local decoder, larger neighborhood, or spatial reader can recover a distinction that was never encoded.

The keyed field moves this failure mode. For an ideal keyed assignment, a candidate pair collides with probability $1/4$ at any particular cell, but a collision at one position does not force a collision at another. Adding an issued candidate leaves every existing fingerprint definition unchanged; the price is the explicit, quantifiable reduction in locally unique anchors given by Eq. (5), rather than a deterministic whole-field alias. GSF does not claim collision-free operation at arbitrary depth. It uses a fixed interface for a changing finite candidate set and exposes where that interface is locally insufficient.

Rejection happens at two levels. At the level of the code, four active labels fill all four two-bit symbols, so a hard readout under the direct code always names a label, whereas keyed candidate matching can return no match or several matches. At the level of the readout rule, either code can abstain on low decoder confidence; the coverage-matched selective rule of Appendix K.3 is such a rule for the direct code. What a confidence threshold cannot supply is the code-level signal: a symbol read with a high margin that matches no issued candidate is, under the keyed code, evidence that the cell no longer carries any issued generation, and the verifier counts such cells as positive OPEN evidence (C_{nokey} in Appendix B); under the direct code the same symbol is another legal label. Table 40 isolates the code-level behavior on carrier-destroyed cells, read as hard symbols without the readability mask that the verifier applies.

Table 40: Local outcome decomposition on carrier-destroyed cells, read as hard symbols without the verifier's readability mask. This is not a total-system ranking: abstention counts as an error in unconditional accuracy, but preserves the distinction between unsupported and asserted history.

Local outcome	Direct lookup	Keyed exact
Correct generation ↑	42.04%	17.60%
Wrong generation ↓	57.96%	27.52%
Abstain / ambiguous	0.00%	54.88%

The direct rule is correct more often in this subset, 42.04% versus 17.60%, when abstention is scored as error. That advantage and its failure mode are two sides of the same compulsory decision: it also asserts a wrong generation on the majority of these cells and never rejects at the code level. For reissuable provenance, these outcomes are not interchangeable. A no-match does not solve the cell, but it leaves an honest uncertainty state for spatial inference or an UNKNOWN decision; a confident wrong generation turns damaged evidence into asserted history. The keyed field's no-match and multi-match outcomes are therefore not merely lost accuracy; they also form an explicit uncertainty channel.

The field encodes compatibility, not a public version number. The direct symbol expresses $a \bmod 4$ and repeats that convention across positions and lineages. The GSF symbol evaluates compatibility for the full tuple (K, L, a, i): changing the key, lineage, generation, or cell changes the expected field. The keyed-permuted direct control shows that a direct lookup can be masked by a key and position, but masking does not make generations a and $a + 4$ distinct. Both still receive the same padded symbol at every cell. This is why a keyed candidate-compatibility field is not equivalent to hiding a two-bit version tag.

K.5 Width Trades Against Spatial Granularity

Increasing the direct code width postpones deterministic aliasing from four to 2^b labels, but it spends the same fixed carrier budget in a different place. With 320 binary slots, $b = 2, 3, 4, 8$ provide 160, 106, 80, and 40 logical regions, respectively. Coarser groups are also more likely to cross a true revision boundary; at depth three, 58.05% of the $b = 8$ groups are mixed. A wider direct tag can therefore buy a longer closed label window only by reducing the resolution of the regional history map. It does not jointly obtain collision-free arbitrary candidate depth, the same two-bit cell rate, and the same spatial granularity. Appendix L gives the full fixed-slot width law, break-even depths, and mixed-group analysis.

K.6 Editing Creates a Different Information Loss

The preceding argument concerns what the representation can express. Editing adds a second failure mode by changing the observed symbol itself. Define a decoded-code-flip cell as one whose hard two-bit observation differs from the symbol written for its true generation. Conditional on that event, neither keyed exact matching nor direct lookup can recover the true label from the damaged cell alone: both are exactly zero in every depth-three edit condition in Table 41.

Table 41: Recovery on decoded-code-flip cells in the five depth-three edit conditions. Cells/image is the mean number of affected cells. Local accuracies are structurally zero; the spatial reader uses evidence outside the damaged cell.

		Generation accuracy ↑		
Query	Cells/image	Keyed exact	Direct lookup	GSF spatial
U_0 at head 3	31.8	0.0000	0.0000	0.2248
U_1 at head 3	32.7	0.0000	0.0000	0.2188
U_2 at head 3	31.0	0.0000	0.0000	0.2355
U_3 at head 3	31.4	0.0000	0.0000	0.2303
U_p at head 3	31.6	0.0000	0.0000	0.2377

The GSF reader remains nonzero, at 0.2188–0.2377, because it can consult surviving temporal anchors and coherent neighboring regions. This experiment establishes a deliberately narrow point: nonlocal evidence is needed after a physical symbol flip even when the local code is collision-free over the active labels. It is not presented as a ranking against every possible direct-code spatial reader. Rather, it shows that removing keyed local collisions would not remove the inference problem created by editing.

K.7 Carrier and Code Are Jointly Optimized

The code-only construction isolates representation, but a codebook is not a free plug-in to the physical system. A complementary control writes each assignment through the same frozen carrier, decoder, and alignment stack on 12 source images and ten post-processing conditions. On the clean generation-zero issue, exact two-bit readback is 1.0000 for the native keyed assignment, 0.9299 for direct assignment, and 0.9053 for keyed-permuted direct assignment. Across the ten conditions, decoded-symbol change rates range from 3.6–10.3% for the native code, 13.3–25.9% for direct recoding, and 13.1–21.9% for keyed-permuted direct recoding.

Table 42 compares physical recoding with the slot-level counterfactual on non-saturated cells. Their agreement is 0.72–0.87 rather than near identity, whereas the largest difference between positive- and negative-written-slot change probabilities is only 0.0095. Thus a simple bit-sign asymmetry does not explain the discrepancy.

Table 42: Agreement between physical recoding and the code-only slot-level counterfactual of Eq. (22) on non-saturated generation slots. The last column is the absolute positive/negative written-sign asymmetry.

Condition	Direct agreement ↑	Permuted-direct agreement ↑	Sign asymmetry ↓
Clean	0.8499	0.8471	0.0058
Blur $\sigma = 1.0$	0.7677	0.8191	0.0037
Brightness 1.15	0.8567	0.8535	0.0047
Brightness 1.30	0.8544	0.8271	0.0069
Contrast 1.50	0.8673	0.8465	0.0088
JPEG 50	0.7247	0.7711	0.0014
JPEG 75	0.8028	0.8300	0.0059
Platform-style pipeline	0.7196	0.8120	0.0036
Resize 256	0.7863	0.8008	0.0095
Resize 384	0.8353	0.8451	0.0093

The coupling has a concrete source: the translation stage of alignment scores candidate translations against the native keyed expected signs, whereas rotation and scale are chosen by the key-independent decoder margin. Replacing the code therefore changes both the semantic lookup after decoding and the translation objective before decoding. The counterfactual should consequently be read as a clean comparison of representation geometry that favors the direct code. The physical control shows only that a complete direct alternative would require joint optimization and fresh evaluation of the writer, decoder, and alignment rule; it is not an intrinsic upper bound on such a system.

K.8 Why the Keyed Operating Point

On a fixed alphabet of four labels the direct code is the better local code, and the price of the keyed field can be stated in numbers: on intact issued cells the code-only counterfactual gives the direct code 0.99983 unconditional accuracy against 0.5788 for keyed exact matching, a gap the spatial reader closes only in part (Table 2), and direct recoding on the same carrier still reads back 0.9299 of its symbols. GSF pays that price for four properties the direct code does not have. When the candidate set grows, the direct code aliases the fifth label with an earlier one at every position, whereas a keyed candidate leaves every existing fingerprint unchanged and costs only the predictable loss of unique anchors of Eq. (5). A readable symbol that matches no issued candidate is evidence under the keyed code and a legal label under the direct code, so the field's rejections feed the verifier instead of being absorbed into a wrong generation. Local ambiguity varies by position, so a spatial reader can recover part of the regional history from other cells, whereas a global alias leaves no such cell. And the frozen writer, decoder, and alignment rule are co-adapted to the keyed code, so a direct alternative is a different system to be optimized and evaluated, not a drop-in replacement. A deployment whose label set never exceeds four could adopt the direct code outright; the keyed field is the choice once the candidate set is allowed to grow.

All representation-level results use the same 480 roots, candidate sets, revision fields, 320 physical slots, and 17 query conditions. Alternative assignments are generated deterministically from the issued label, cell coordinate, lineage key, and fixed slot evidence. Reinstating the native keyed assignment reproduces every original local and spatial input, and the two-bit rows independently match the fixed-slot width experiment. These reconstruction checks keep the two claims separate: the direct code's four-label advantage is measured rather than assumed, and the keyed field is selected for the representation that a reissuable field requires.

L A Fixed-Slot Design Law for Temporal Fingerprints

Wider fingerprints reduce the probability that two candidates predict the same local symbol, but a fixed carrier cannot increase width without reducing the number of spatial regions. We derive the resulting anchor law, verify it at five widths, and show that two bits provide the strongest worst-case balance over the three evaluated successor depths. The same law also predicts how a deeper lineage gradually shifts the preferred width.

L.1 A Fixed-Slot Anchor Lemma

Let a writer provide S_{slot} physical binary generation slots. Grouping b consecutive slots gives $M_b = \lfloor S_{\text{slot}}/b \rfloor$ logical regions. Assume that, at each logical region, the b-bit fingerprints of different candidates are independent and uniform. If the true label must be separated from d other candidates, a region is a unique temporal anchor with probability $(1-2^{-b})^d$. Linearity of expectation therefore gives

$$N_S(b,d) = \left\lfloor \frac{S_{\text{slot}}}{b} \right\rfloor (1-2^{-b})^d. \tag{24}$$

For a declared depth set D, comparison to the best width separately at each depth yields

$$\eta_D(b) = \min_{d\in D} \frac{N_S(b,d)}{\max_{k\geq 1} N_S(k,d)}. \tag{25}$$

Code width is frozen when the carrier is instantiated, whereas the realized candidate depth within the declared range D is not known at that point. A single width must therefore serve every $d \in D$. Normalizing by the best allocation at each depth places these operating points on a common relative-efficiency scale, and the outer minimum selects the width that preserves the largest worst-case fraction of the depth-wise optimum. This criterion follows from the ex-ante frozen-width decision and is independent of the reader outcomes reported below.

Normalized-maximin result. Ignoring only the finite grouping remainder, the common factor S_{slot} cancels and the comparison becomes $b^{-1}(1-2^{-b})^d$. At $D = \{1,2,3\}$, the per-depth maximizing widths are $1, 2, 3$. Width two has relative efficiency $3/4$ at $d = 1$, one at $d = 2$, and greater than $3/4$ at $d = 3$. Width one falls to $192/343 \approx 0.560$ at $d = 3$. For every integer $b \geq 3$, its relative efficiency at $d = 1$ is at most

$$\frac{2(1-2^{-b})}{b} \leq \frac{7}{12} < \frac{3}{4}. \tag{26}$$

Thus $b = 2$ is the unique normalized-maximin allocation under the fixed-slot, keyed-iid assumptions and this declared depth range.

Our writer has $S_{\text{slot}} = 320$. The main slot order lists the 160 generation cells in row-major grid order and places their two channel slots consecutively; width-b groups are consecutive blocks in this 320-slot list. The $b = 3$ case discards the final two slots, the largest remainder among the tested widths: $2/320 = 0.625\%$. Table 43 shows that finite grouping does not change the winner or any relevant ordering.

Table 43: Expected unique anchors per image from Eq. (24). The final row is finite-S normalized maximin efficiency over $D = \{1,2,3\}$.

	Expected unique anchors ↑				
d	$b=1$	$b=2$	$b=3$	$b=4$	$b=8$
1	160.00	120.00	92.75	75.00	39.84
2	80.00	90.00	81.16	70.31	39.69
3	40.00	67.50	71.01	65.92	39.53
4	20.00	50.62	62.14	61.80	39.38
5	10.00	37.97	54.37	57.94	39.22
6	5.00	28.48	47.57	54.31	39.07
7	2.50	21.36	41.63	50.92	38.92
$\eta_{\{1,2,3\}}(b)\uparrow$	0.5633	**0.7500**	0.5797	0.4688	0.2490

Figure 11 makes the trade-off visible: no width dominates at every depth, but two bits preserve the strongest worst-case fraction of the depth-wise optimum over the evaluated range.

L.2 When Does an Additional Bit Pay for Itself?

In the continuous-slot comparison, the first depth at which width $b+1$ provides more total anchors than width b is

$$d_b^* = \min\left\{ d \in \mathbb{N} : \frac{(1-2^{-(b+1)})^d}{b+1} > \frac{(1-2^{-b})^d}{b} \right\}. \tag{27}$$

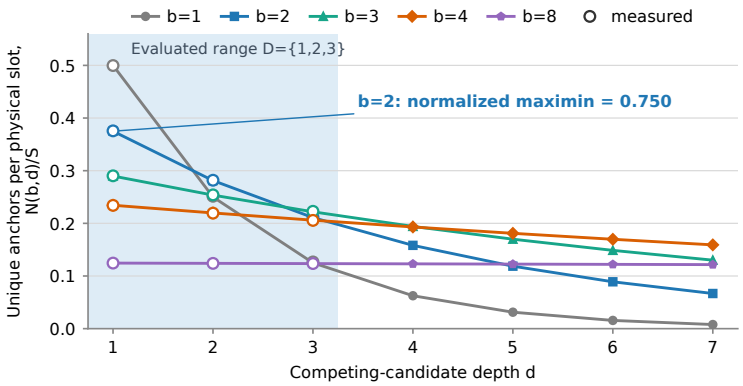

Figure 11: Fixed-slot anchor yield across candidate depth. Curves show $N_S(b,d)/S_{\text{slot}}$; hollow markers are measurements at the three evaluated depths. The shaded region is $D = \{1,2,3\}$, where $\eta_D(2) = 0.750$ is the worst relative efficiency, not the plotted anchor yield.

The full ladder is shown in Table 44. The second bit already pays for itself once there are two competing candidates; the third at three. Further bits pay progressively later. In particular, the eighth bit does not beat the seventh until $d = 34$. The preferred width therefore moves gradually with lineage depth rather than being a depth-independent constant.

Table 44: First competing-candidate depth at which each additional bit increases total unique-anchor count under a fixed continuous slot budget.

Added bit	2nd	3rd	4th	5th	6th	7th	8th
Break-even depth d_b^*	2	3	5	7	12	20	34

L.3 The Collision Identity Holds Across Widths

For keyed fingerprints the blind fraction is $1-(1-2^{-b})^d$, so the uniquely readable fraction is $(1-2^{-b})^d$. Table 45 evaluates all 15 width–depth points using intact issued fields. The largest absolute theory–measurement difference is 0.002857, at $b = 3, d = 3$. Bit width therefore buys collision resistance exactly as predicted; the design question is how many spatial anchors remain after that width is charged against a fixed slot budget.

Table 45: Closed-form and observed uniquely readable fractions for keyed fingerprints. Observations pool logical regions across 480 roots; values in the main Results table are instead per-image medians.

	$d=1$		$d=2$		$d=3$	
b	Theory ↑	Observed ↑	Theory ↑	Observed ↑	Theory ↑	Observed ↑
1	0.5000	0.4997	0.2500	0.2514	0.1250	0.1267
2	0.7500	0.7513	0.5625	0.5637	0.4219	0.4234
3	0.8750	0.8763	0.7656	0.7670	0.6699	0.6728
4	0.9375	0.9365	0.8789	0.8775	0.8240	0.8224
8	0.9961	0.9962	0.9922	0.9918	0.9883	0.9892

L.4 Width Is Invisible to the Physical-Slot Law

Let $s_{a,j}^{(b)}$ be the expected sign of candidate a at physical slot j under width b, and let the true candidate be a^*. The truth-aligned slot evidence is $E_j = s_{a^*,j}^{(b)}\ell_j$. Under iid keyed codewords and the fixed-slot counterfactual of Eq. (22), the distribution of E_j is independent of b: changing width neither adds observations nor changes their marginal signed-evidence law. Width changes which slots form a logical region, how often another candidate matches the entire group, and how that region intersects a revision boundary.

We test this invariance with independently trained keyed two- and eight-bit readers under the same frozen data and optimization rules. As Table 46 shows, the eight-bit reader is numerically higher

in all three reported domains: by 0.0060 over all valid generation slots and by 0.0250–0.0288 on the two smaller damaged subdomains. Because the readers were trained independently, these values provide a descriptive reader comparison rather than a causal decomposition of the difference. The physical-slot result above has a narrower scope: it concerns the marginal truth-aligned evidence law, not equality between independently trained and differently grouped readers. Table 46 therefore pairs the observed reader differences with the spatial product delivered under the same physical-slot budget.

Table 46: Local reader accuracy and regional-history cost for separately trained keyed readers under the same 320-slot budget. "All" is the pooled 17-condition generation-slot domain; "decoded flip" and "carrier-destroyed" are nested diagnostic domains.

Metric	Keyed 2-bit	Keyed 8-bit	Eight-bit change
Conditional reader accuracy			
All valid generation slots, accuracy ↑	0.8539	0.8599	$+0.0060$
Decoded-code-flip slots, accuracy ↑	0.2743	0.3031	$+0.0288$
Carrier-destroyed slots, accuracy ↑	0.4685	0.4935	$+0.0250$
Regional-history product under the same 320-slot budget			
Logical regions ↑	160	40	one quarter as many
Depth-3 mixed groups ↓	0.0000	0.5805	$+58.05$ percentage points
$\eta_{\{1,2,3\}}(b)$ ↑	0.7500	0.2490	-0.5010

The eight-bit reader therefore obtains an observed local conditional advantage of at most 0.0288. Under the same 320-slot budget, that advantage comes with one quarter as many logical regions—40 rather than 160—while 58.05% of depth-three groups cross a generation boundary and normalized-maximin efficiency is 0.249 rather than 0.750. Eight bits improve a local readout statistic by producing a substantially coarser spatial product; they do not dominate the two-bit allocation on the same regional-history objective.

L.5 Fully Damaged Windows Contain No Clean Local Evidence

A fully damaged logical window contains no constituent physical slot whose hard observation retains its written symbol. Head-zero queries contain only one candidate, so correctness there is vacuous and is excluded from this diagnostic. For heads with at least one competing candidate, keyed and direct one- and two-bit local decisions are exactly zero over 160,101 one-bit windows and 79,312 two-bit windows. No clean local evidence in those windows points to the original code.

Widths three and above show residual local accuracy of 0.024–0.048, but this does not reverse the information statement. Their coarser groups can cross a ground-truth revision boundary; a minority subset of slots may then agree with the label used to score the mixed group by chance. The nonzero value is a boundary-mixing effect, not recovery of an erased homogeneous codeword.

L.6 Coarse Grouping Changes the Spatial Product

Table 47 expands the boundary-mixing row of Table 46 across all five widths and three issued successor depths. It measures the fraction of logical groups that contain physical slots from more than one ground-truth generation. One- and two-bit groups preserve the original cell support. At eight bits, 58.05% of groups in the deepest field cross a generation boundary. This statistic measures grouping granularity rather than pixelwise localization error.

The main grouping follows row-major generation-cell order. Repeating the analysis with column-major order changes aggregate keyed accuracy by 0.0027 at $b = 4$ but by 0.0194 at $b = 8$. Coarse groups are consequently more sensitive to an otherwise arbitrary grouping geometry. Together with the normalized anchor efficiencies in Table 43, this explains the three spatial costs of very wide codes: fewer logical regions, more mixed regions, and greater grouping sensitivity.

L.7 Keyed and Direct Codes Allocate Ambiguity Differently

The keyed lemma describes a variable-candidate fingerprint family. A position-invariant direct code instead eliminates stochastic collisions inside its finite injective window and aliases deterministically

Table 47: Fraction of logical groups crossing a ground-truth generation boundary at each issued successor depth.

	Mixed-group fraction ↓			
Width	Depth 1	Depth 2	Depth 3	Logical regions ↑
1	0.0000	0.0000	0.0000	320
2	0.0000	0.0000	0.0000	160
3	0.1961	0.2589	0.2808	106
4	0.1871	0.2459	0.2661	80
8	0.4547	0.5488	0.5805	40

when the alphabet is exhausted. Appendix K gives the matched two-bit comparison and shows that either family still loses local information after a decoded-symbol change. At one bit, the keyed field is ambiguous on 87.5% of regions at $d = 3$, while the direct field already aliases labels 0 with 2 and 1 with 3.

L.8 Local Aggregation and Spatial Cost Move Together

Wider groups can aggregate more weak local evidence, and direct assignments can sharpen decisions inside a finite injective window. Under a fixed carrier, however, those gains arrive together with the spatial costs measured above. At eight bits the field contains only 40 logical regions, 58.05% of deepest groups cross a generation boundary, and normalized maximin efficiency falls to 0.249 over the evaluated depth range. The design law therefore compares collision resistance and spatial evidence jointly rather than optimizing a single local readout statistic.

L.9 Complete Results and Reproducibility

The full width-by-readout matrices cover all 17 query conditions for keyed and direct codes at $b \in \{1, 2, 3, 4, 8\}$. The accompanying artifacts provide these matrices in machine-readable form together with the code that reproduces the anchor counts, collision identities, grouping statistics, and decision rules reported above.

Validation. Keyed width two uses the native GSF code; the other widths use disjoint PRF domains over the same 320 ordered slots. Independent executions agree on all 8,160 width-two query rows. The theoretical and empirical results therefore close the same loop: wider fingerprints reduce local collision by spending spatial regions, while two bits uniquely maximize worst-case relative anchor efficiency over the evaluated depths.

M Transferring the Reissuance Protocol to Released Watermarks

Section 5.5 shows that the reissuance protocol transfers to released watermarks (Lu et al., 2025; Bui et al., 2025; Zhang et al., 2025; Sander et al., 2025) and asks which parts of the joint task transfer with it. This appendix states what is shared, what is method-specific, and what each number in Table 3 and Figure 4(e) means. All results are on the 480 roots of Section 5.1.

M.1 A Common Task with Method-Specific Constructions

A released watermark has no vocabulary for reissuance. In a separate 24-root check with released decoders and a constant message, under inpainting and super-resolution edits rather than the editors of Section 5.1, the current issue and its archived predecessor both read "present" on every root for all four marks tested (VINE-R, TrustMark-Q, EditGuard, WAM), with no difference that could express SUPERSEDED, and edited outputs read "present" on 62.5–100% of roots depending on the mark. The comparison therefore gives each released watermark the reissuance protocol of Section 4, and asks which parts of the task it can then support. The transferred carriers are post-hoc marks in the encoder–decoder tradition (Zhu et al., 2018); in-generation watermarks (Fernandez et al., 2023; Wen et al., 2023; Yang et al., 2024) embed at sampling time and cannot be written into an already rendered image, so they are not transferred.

Shared task. Every system follows the same 480 roots through four issues I_0–I_3 under the same edit-instruction table: the 2,400 rows of backbone, prompt, seed, operation, and parent that produced GSF's outputs are matched row by row. Each system generates its own lineage from the source image, so its issued and unissued pixels differ from GSF's while the instructions coincide; this costs 12,000 editor calls. The queries per system are 4×480 under an issue's own head (truth CURRENT), 6×480 under a later head (truth SUPERSEDED), and six unissued query conditions (truth OPEN): the outputs before issues 1–3 at their own heads, the output after I_3 at head 3, and the branch at its parent's head and at head 3. The seventh condition of Table 16, a branch from I_1 queried at head 2, exists only for GSF and is excluded here. Regional history is read on the same current issues at depths 1–3.

Shared protocol. For the global-message watermarks every reissue writes a generation-conditioned 100-bit issue message: a lineage identifier and a keyed content descriptor, protected by the same error-correcting format and masked by a keystream conditioned on lineage and generation. The live head and the registry are those of Section 4. At query time every system is decided by the routing chain of Table 10 at the operating point of the main results, applied to the evidence its carrier provides. For a global message that evidence is the message alone: a decoded message that names the head is CURRENT, a decoded predecessor is SUPERSEDED, and an undecodable message is OPEN, the route that also carries three quarters of GSF's own OPEN decisions (Table 18); with no field to lose, the chain's abstention route never fires. WAM's adapted field takes the field routes. The issuance semantics and live-head rule are shared; each construction maps its available evidence to the task outputs as specified in Table 48.[1] No baseline network is retrained. What transfers with the protocol is the issue decision against the live head; what the comparison then measures is how much of the remaining task each carrier's edit evidence supports.

Table 48: Method-specific construction of the compared systems. A constant field is a deliberately non-spatial reference: it labels every cell with the decoded current generation and cannot recover inherited regional history.

Watermark	Reissuance	Evidence for an unissued output	Regional-history readout
VINE-R	Rewrite the generation-conditioned issue message over the complete image.	Message survival yields an issue match; message loss yields OPEN.	Head label at every cell.
TrustMark-Q	Rewrite the same issue message with the released quality-oriented TrustMark model.	Message survival yields an issue match; message loss yields OPEN.	Head label at every cell.
TrustMark-RM	Apply the released remover before rewriting the next generation-conditioned TrustMark message.	The rewritten message has the same two outcomes.	Head label at every cell.
OmniGuard	Rewrite the issue message with the released robust message watermark.	The released tamper-mask output supports OPEN; message loss yields OPEN.	Head label at every cell; the snapshot mask carries no earlier generations.
WAM	A contiguous generation block and a 160-cell field in GSF's layout; reissue rewrites the block and the committed cells.	An unreadable generation block or a disrupted cluster supports OPEN; a field lost entirely or disrupted beyond one cluster abstains.	Generation code decoded on cells with readable WAM evidence.
GSF	Lineage evidence, generation-conditioned issue evidence, live head, and inherited keyed field.	Once the image is attributed to the lineage, undecodable issue evidence or a disrupted cluster supports OPEN without asserting an issue.	Spatial reader recovers the last authorized generation at every field cell.

Readability on clean images. On the clean images at head 0, the issue message is recovered on 479/480 roots for VINE-R and on 480/480 for both TrustMark variants; OmniGuard recovers it on 480/480 but reads 30 clean issues as OPEN because its own tamper mask reports tampering (90 such cases over the four issues at their own heads). WAM's generation block, the contiguous cells that carry its generation code, is readable on 442/480 clean roots and 391 read CURRENT. No root is excluded from any reported number; the WAM row is a diagnostic of a GSF-like layout realized through WAM, whose native localized-message task is a separate product.

WAM's layout. WAM writes one message per masked region and composes the regions in the frame, whereas GSF's field is addressable per cell (Eq. 7), so one write carries every inherited

[1] TrustMark-RM applies the released remover, which the ICCV version of TrustMark names ReMark (Bui et al., 2025); we keep the released-code name.

label at once. Coexisting WAM messages interfere: with three (lineage header, generation block, field), the header decodes on only 185/480 clean roots; with one message per field cell, every cell is detected as marked yet none of the 160 returns its own code. The adapted layout therefore keeps coexisting messages few: it writes the generation code as one block and rewrites only the committed cells at each reissue, leaving inherited cells as previously written; Table 52 measures what survives of them.

M.2 CURRENCY TRANSFERS

Table 49 expands the issued columns of Table 3 by live head. With a generation-conditioned message and a live head, VINE-R, both TrustMark variants, and OmniGuard separate CURRENT from SUPERSEDED at every head, and the two TrustMark variants differ by at most 0.4 points under any condition. Currency transfers with the protocol wherever the carrier's rewritten message stays readable, and the released robust watermarks support it as accurately as GSF or slightly better. WAM is the exception, and its failure has a structure. It reads its generation block on the I_0 images (81.5% CURRENT at head 0) but loses issue identity once a real edit has been reissued (0.2%, 0.0%, and 0.0% at heads 1–3, where repeated editing and re-embedding leave its field disrupted and the shared chain returns OPEN or abstains); its 40.8% SUPERSEDED is the same set of readable I_0 images correctly expired under later heads, not a separate ability.

Table 49: Issued images by live head (% CURRENT/SUPERSEDED accuracy). At head h the CURRENT condition has 480 queries and the SUPERSEDED conditions $480h$; overall columns contain 4×480 and 6×480 queries. "+" as in Table 3; the GSF row is Table 15.

	CURRENT/SUPERSEDED accuracy (%) ↑				
System	$h = 0$	$h = 1$	$h = 2$	$h = 3$	Overall
+ VINE-R	99.8/–	98.3/99.8	97.1/99.1	97.7/98.4	98.2/98.9
+ TrustMark-Q	**100.0**/–	**100.0/100.0**	**100.0/100.0**	**99.8/100.0**	**99.9/100.0**
+ TrustMark-RM	**100.0**/–	99.6/100.0	99.6/99.8	99.4/99.7	99.6/99.8
+ OmniGuard + tamper mask	93.8/–	95.2/93.8	96.5/94.5	95.8/95.1	95.3/94.7
+ WAM in GSF layout[†]	81.5/–	0.2/81.5	0.0/40.8	0.0/27.2	20.4/40.8
Full GSF (ours)	99.6/–	99.8/99.6	99.8/99.8	99.0/99.8	99.5/99.8

M.3 UNISSUED RECOGNITION NEEDS EVIDENCE BEYOND THE MESSAGE

Table 50 gives the six unissued query conditions with 95% Clopper–Pearson intervals; the GSF row is the sum of the corresponding six rows of Table 16. For the three message-only watermarks (VINE-R, TrustMark-Q, TrustMark-RM) the editor decides the outcome. A destroyed message reads as OPEN: 63.5%, 70.1%, and 70.5%. A surviving message names the issue it was written for, which is a false CURRENT at that issue's own head and a false SUPERSEDED after the head advances (VINE-R 879 and 172 of its 1,051 false issues; TrustMark-Q 722 and 140; TrustMark-RM 710 and 139). SDXL-Turbo destroys every message of every watermark (0 of 730 readable), FLUX.1 Kontext preserves the most (521 of 720 for VINE-R), with InstructPix2Pix and Qwen Image Edit in between. An unissued output whose message survives thus reads as the issue it names; GSF's field adds a second observation, a disrupted cluster, which catches most such outputs and halves false acceptance (Tables 17 and 18). On these conditions GSF's 94.0% is 75.1 points of undecodable issue evidence, the route it shares with the message-only watermarks, and 18.9 points of field evidence, 17.1 of them from a disrupted cluster. GSF's own recall on the four sequential conditions is 96.04%, 94.17%, 92.92%, and 91.67% (outputs before issues 1–3 and after I_3), a slow decline with depth.

OmniGuard's 61.3% splits the same way: 35.0 points of undecodable message and 26.3 from its tamper mask, a second observation; its robust message nevertheless survives more edits than any other watermark, so it also has the highest false-issue rate, 38.7%. WAM's 0.4% false issue must be read with its 48.8% UNKNOWN: after the first real edit its generation block is unreadable, and its field is lost entirely on 1,164 of the 2,880 outputs and disrupted beyond one cluster on 242 more, the two routes that abstain under the shared chain; its 50.8% OPEN is decoding failure rather than discrimination.

Table 50: Unissued outputs on the six query conditions shared by all systems ($6 \times 480 = 2{,}880$ queries; % with 95% Clopper–Pearson intervals). False issue is a prediction of CURRENT or SUPERSEDED; UNKNOWN is never credited as OPEN; –: no abstention route.

System	OPEN ↑	False issue ↓	UNKNOWN ↓
+ VINE-R	63.5 [61.7, 65.3]	36.5 [34.7, 38.3]	–
+ TrustMark-Q	70.1 [68.4, 71.7]	29.9 [28.3, 31.6]	–
+ TrustMark-RM	70.5 [68.8, 72.2]	29.5 [27.8, 31.2]	–
+ OmniGuard + tamper mask	61.3 [59.5, 63.1]	38.7 [36.9, 40.5]	–
+ WAM in GSF layout†	50.8 [48.9, 52.6]	0.4 [0.2, 0.7]	48.8 [47.0, 50.7]
Full GSF (ours)	**94.0 [93.0, 94.8]**	2.1 [1.6, 2.7]	3.9 [3.2, 4.7]

M.4 A CONTENT COMMITMENT TRADES CURRENCY FOR CAUTION

The message's content descriptor does separate unissued outputs from issued images: as a scalar distance it reaches an AUROC of 0.92–0.94 for every watermark (0.945 for GSF's own commitment). Thresholding it, however, only moves outcomes between CURRENT and withheld (Table 51). At the strict threshold $\tau_{\mathrm{id}} = 1/48$, CURRENT accuracy under an issue's own head falls to 67.4–79.9% while OPEN recall rises by only 1.8–5.3 points for the message-only watermarks and 4.5 for OmniGuard; false issue drops to 7.1–7.8% because surviving messages are now withheld rather than accepted. Relaxing to $2/48$ restores 90.1–97.6% CURRENT at the same OPEN recall. At the thresholds and routing tested, a content commitment trades currency for caution along one scalar and adds little OPEN recall.

Table 51: Effect of a strict content commitment on the global-message watermarks (%). Native: the protocol as evaluated in Table 3; τ_{id}: the message's content-descriptor mismatch threshold in units of $1/48$. Own/later: CURRENT/SUPERSEDED accuracy under the issue's own head and a later head; OPEN and false issue on the six unissued query conditions.

System	Native own / later / Open / false issue	$\tau_{\mathrm{id}} = 1/48$ own / later / Open / false issue	$\tau_{\mathrm{id}} = 2/48$ own / later / Open / false issue
+ VINE-R	98.2 / 98.9 / 63.5 / 36.5	75.7 / 98.9 / 68.9 / 7.8	94.2 / 98.9 / 68.9 / 10.8
+ TrustMark-Q	99.9 / 100.0 / 70.1 / 29.9	79.8 / 100.0 / 72.0 / 7.1	97.6 / 100.0 / 72.0 / 10.6
+ TrustMark-RM	99.6 / 99.8 / 70.5 / 29.5	79.9 / 99.8 / 72.4 / 7.1	97.6 / 99.8 / 72.4 / 10.2
+ OmniGuard + tamper mask	95.3 / 94.7 / 61.3 / 38.7	67.4 / 94.7 / 65.8 / 7.7	90.1 / 94.7 / 65.8 / 11.3

M.5 REGIONAL HISTORY

Table 52 reports the map metrics of Table 2 for every system, adding the newest-slice IoU used by Figure 4(e). A global message produces no field, so its map is a constant field: every cell labeled with the decoded head. Its accuracy is the area of the latest committed scope and its newest-slice IoU coincides with that accuracy; its mIoU falls with depth because every inherited generation is missed. The constant field was scored over all 256 cells of the grid, whereas WAM and GSF are scored on the 160 field cells; on GSF's own fields it gives 27.50%, 18.12%, and 14.30% mIoU on the 160 cells against 28.71%, 19.14%, and 14.89% on all 256, so the difference is about one point and does not affect any comparison. WAM's adapted field is the closest spatial counterpart. On the cells where its evidence survives it is accurate (91.3–94.7% cell accuracy), but 26.9–36.9% of cells are unreadable after repeated editing and re-embedding and count as errors, so its mIoU falls from 49.05% to 24.30%; its high newest-slice IoU reflects that the most recently rewritten cells are exactly the ones still readable. Localizing a single edit is not where the systems differ: on unissued outputs, the average precision of OmniGuard's mask, WAM's evidence map, and GSF's raw disruption map against the proposed scope is 0.625, 0.642, and 0.634. The separation lies in inherited history, which only the keyed field carries.

M.6 DELIVERY RANGE

Table 53 reports CURRENT retention of the current issue at head 3 under the 22 conditions of Appendix E that were run for every system (the ten finer levels of that appendix were run for GSF only). The message-only watermarks hold 95.6–98.1% (VINE-R), 84.4–99.8% (TrustMark-Q), and 81.2–99.4% (TrustMark-RM) across all 22 conditions, OmniGuard 52.9–96.2%, whereas GSF falls to 0–65% under JPEG at quality 50 or below, resizing to 192 pixels, the platform pipelines, contrast gain, and brightness $\times 1.30$, and to 72.5% at brightness $\times 1.15$. WAM is omitted: its current issue reads CURRENT in 0.0% of clean roots at head 3 (Table 49), so delivery adds nothing. A robust static

message thus survives a wider range of delivery transformations; the same survival keeps naming the old issue on an unissued output and carries no history. Over the same 22 conditions, GSF's median map accuracy stays at or above 85.00% (Table 20).

Table 52: Regional history at depths 1–3 (median % over 480 roots). Acc. and mIoU I_3 at head 3 as in Table 2; Newest IoU is the IoU of the newest slice with the committed scope. Global-message rows are the constant field scored over all 256 cells (see text); WAM and GSF are scored on the 160 field cells. Baseline rows carry the protocol of Table 48.

Depth	System	Acc. ↑	mIoU ↑	Newest IoU ↑
1	+ VINE-R	57.81	28.91	57.81
1	+ TrustMark-Q	57.03	28.52	57.03
1	+ TrustMark-RM	57.03	28.52	57.03
1	+ OmniGuard	56.64	28.32	56.64
1	+ WAM†	60.00	49.05	90.97
1	**GSF (ours)**	**97.19**	**93.87**	**94.92**
2	+ VINE-R	57.81	19.40	57.81
2	+ TrustMark-Q	57.81	19.27	57.81
2	+ TrustMark-RM	57.42	19.21	57.42
2	+ OmniGuard	57.03	19.08	57.03
2	+ WAM†	52.50	32.38	93.92
2	**GSF (ours)**	**95.00**	**87.27**	**94.09**
3	+ VINE-R	58.20	14.70	58.20
3	+ TrustMark-Q	57.03	14.40	57.03
3	+ TrustMark-RM	57.42	14.65	57.42
3	+ OmniGuard	58.20	14.65	58.20
3	+ WAM†	53.75	24.30	93.44
3	**GSF (ours)**	**93.13**	**78.96**	**92.86**

Table 53: CURRENT retention of the current issue (% of 480 roots). Operators as in Appendix E.1; the GSF column is the I_3 column of Table 19; baseline columns carry the protocol (TM: TrustMark).

Condition	+VINE-R	+TM-Q	+TM-RM	+OmniGuard	GSF
Clean	97.7	99.8	99.4	95.8	99.0
JPEG 75	97.5	99.8	98.8	87.1	95.6
JPEG 50	96.9	98.5	97.5	89.2	35.0
JPEG 40	96.7	97.3	96.9	84.0	1.7
JPEG 30	96.5	94.6	93.8	70.0	0.0
Double JPEG 75	97.7	99.8	98.5	89.0	95.2
Resize 384	97.9	99.8	99.4	83.5	99.4
Resize 320	97.9	99.8	99.4	79.0	99.0
Resize 256	97.7	99.8	99.4	76.0	99.4
Resize 192	97.1	99.8	99.4	72.7	1.0
Blur 0.5	97.9	99.8	99.4	89.6	99.0
Blur 1.0	98.1	99.6	99.4	57.7	98.1
Brightness $\times 0.85$	97.7	99.8	99.4	94.0	99.4
Brightness $\times 1.15$	97.3	98.5	98.3	77.3	72.5
Brightness $\times 1.30$	95.6	94.0	94.2	63.5	39.4
Contrast $\times 0.80$	97.3	99.8	99.4	95.0	99.2
Contrast $\times 1.25$	97.3	93.3	93.3	71.7	64.6
Contrast $\times 1.50$	96.7	84.4	81.2	52.9	31.3
Gamma 0.80	97.9	99.8	99.4	96.2	75.0
Gamma 1.25	96.7	99.6	99.0	77.7	66.7
Platform A	97.1	98.8	97.5	92.3	58.1
Platform B	97.1	99.4	98.3	92.7	2.5

M.7 A REGISTRY LOOKUP REPRODUCES THE MAP OF AN ISSUED IMAGE

The registry of Section 3 records every committed scope A_g and parent link, so Eq. (1) reconstructs τ_g exactly from a decoded label. A deployment that consults the registry at every verification therefore needs no field on issued images: a robust global message carrying the protocol returns the exact map whenever its label decodes, at the carrier's status accuracy—for TrustMark-Q, 99.9/100.0 on clean images and 84.4–99.8% CURRENT across the 22 delivery conditions of Table 53—on clean images above the 93.13% that GSF reads from the pixels at depth 3. That design keeps the history in the registry and reads it online. GSF keeps it in the copy (Appendix J.3): a partner entrusted with K and the list of issues reads each region's last authorized revision without the registry's records, and an unissued output, on which a surviving message names the old issue (Table 50), shows in its field where it departs from its parent (Figure 3). Both designs run the same reissuance protocol—generation-conditioned issue message, live head, registry, and reissuance transaction—and it is the protocol that gives every row of Table 3 supersession without retraining.

M.8 REPRODUCIBILITY NOTES

The five baseline runs (480 lineages each, generated once per system) took about 124 GPU-hours in total; GSF's own run of the 480 lineages took about 23 GPU-hours including the final map decoding. Qwen Image Edit is not deterministic across runs: regenerating six roots changed 91 of 240 editor outputs at the bit level and 36 of 570 verifier decisions, and every divergence originated at a Qwen step. We did not force deterministic kernels, which would change the numerics; comparability rests on the matched instruction table rather than on pixel identity, and GSF's outputs were generated once under the same editors. The verifier is deterministic: the delivery run of Appendix E reproduced all 1,920 earlier clean decisions.

N NEGATIVE QUERIES: IMAGES OUTSIDE THE NAMED LINEAGE

Every query of Section 5.1 presents an image of the lineage it names, so the verifier there decides among the four statuses of an attributed image. This appendix tests the attribution itself: the query names a lineage, the image is unmarked or issued by another lineage, and the verifier evaluates the lineage term λ_L of Eq. (2) before any other evidence, so the correct answer is UNKNOWN. The negatives are a proxy for the open world constructed for this test: three classes of clean image at one head, chosen to exercise the attribution rule. A deployed verifier meets the same question on a wider input—edited and degraded images (Zhao et al., 2024; Saberi et al., 2024), forged marks (Müller et al., 2025), and keys and lineages beyond this study.

Read with its own key at head 3, every I_3 of the evaluation decodes its lineage identifier and its issue label, 3 (480/480 each), and receives the decisions of Table 15. Read with the keys of nine other roots each, none of the 4,320 pairs decodes either, and unmarked images decode neither: 0/480 on COCO val2017 images disjoint from the roots under one root's key each, and 0/480 on the roots' own unmarked sources under their own keys. Over these 5,280 reads the per-candidate false-match rate is at most 0.07% (95% Clopper–Pearson); analytically, a non-member read passes the identifier's BCH check on 4–7% of decodes and must then match 36 bits, about $6\% \times 2^{-36} < 10^{-12}$ per candidate. Association is thus supported for the lineage's own images alone; every image outside it returns UNKNOWN, and none is accepted as an issue (Table 54).

Table 54: Queries outside the named lineage on clean images at head 3 (truth UNKNOWN) and the lineage's own I_3 as control (truth CURRENT): the number of queries whose issue evidence decodes and whose lineage identifier matches, and the decision taken with the lineage term evaluated first. Wrong keys are nine per root; control decisions are those of the main evaluation.

Query	n	Issue evidence	Identifier	Decision
Unmarked image, other content	480	0	0	UNKNOWN 480
Unmarked source of the named lineage	480	0	0	UNKNOWN 480
Issued I_3 of another lineage	4,320	0	0	UNKNOWN 4,320
I_3 of the named lineage (control)	480	480	480	CURRENT 475, OPEN 2, UNK. 3