Compositional Non-Face Re-Identification Pressure under Cumulative Vision Releases

Auditing the "leaky bucket" effect in vision benchmarks

Tirth Joshi • **Honggang Wang***
Katz School of Science and Health, Yeshiva University, New York, USA
tjoshi1@mail.yu.edu honggang.wang@yu.edu

ECCV 2026 | MALMÖ | SEPT 8–12

Katz School of Science and Health

Main message. Removing faces does not make identity risk "reset" with each new release. Dataset expansions, checkpoints, embeddings, indices, and metadata accumulate as side information. We define $vRPI_\alpha$, prove that true-posterior pressure is non-decreasing under added releases, and show large empirical pressure growth on face-masked Re-ID benchmarks.

1. Why Single-Release Privacy Is the Wrong Unit

Vision privacy evaluations usually inspect one release in isolation. Real ecosystems publish artifacts sequentially: data, checkpoints, embeddings, indices, sometimes metadata.

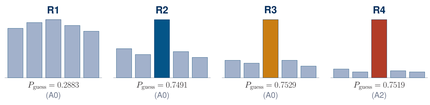
Latent identity. Let $U \in \{1, \dots, N\}$ index identity and X be a raw observation. The non-face transform

$$\bar{X} = \tau(X)$$

removes faces, but identity remains in clothing, gait, and scene context.

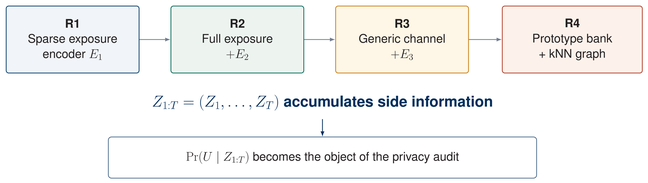
The compositional question. Does the posterior $\Pr(U \mid Z_{1:T})$ sharpen as releases accumulate? The operational target is Bayes-optimal guessing:

$$P_{\text{guess}}(U \mid Z_{1:T}) = \mathbb{E}\big[\max_u \Pr(U = u \mid Z_{1:T})\big].$$



Sketch of $\Pr(U \mid Z_{1:t})$ concentrating across releases – bars are illustrative, not measured. Measured P_{guess} is printed below each panel (Table, §4).

2. Release-as-Channel Formulation



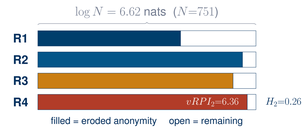
Each disclosure is modeled as a channel $Z_t \sim K_t(\cdot \mid \bar{X})$. The framework is release-sequence aware rather than tied to one Re-ID model or anonymizer.

3. $vRPI_\alpha$: Metric and Core Theory

Rényi pressure index. For $\alpha > 1$, using Arimoto conditional Rényi entropy,

$$\boxed{vRPI_\alpha(T) = \log N - H_\alpha(U \mid Z_{1:T})}.$$

Under a uniform prior, $\log N$ is maximal identity uncertainty: $vRPI_\alpha$ is the anonymity already eroded by cumulative releases.



Rows R1–R4: filled = eroded $vRPI_2$ (R4 uses attacker A2, others A0); open = remaining $H_2(U \mid Z_{1:T})$. Nats, $N = 751$.

Theorem 1: monotonicity on the true posterior

$$H_\alpha(U \mid Z, W) \le H_\alpha(U \mid Z) \Longrightarrow vRPI_\alpha(T+1) \ge vRPI_\alpha(T).$$

Added side information cannot restore identity uncertainty.

Theorem 2: operational guessing bridge

$$P_{\text{guess}}(U \mid Z) \le \exp\Big(\frac{1-\alpha}{\alpha} H_\alpha(U \mid Z)\Big).$$

Thus higher pressure corresponds to a smaller effective anonymity set and stronger identity inference.

Choosing α. $\alpha = 2$ is the recommended stable default; $\alpha = 5$ or 10 emphasizes high-confidence / worst-case posterior concentration and should be paired with calibration diagnostics.

Additional guarantees.

- **Ideal additivity benchmark:** under conditional independence given identity, Sibson information factorizes into an additive upper bound on pressure across releases.
- **Ablation monotonicity:** post-processing that removes information cannot increase measured pressure.
- **Plug-in stability:** pressure estimated from an approximate posterior changes continuously, with an explicit logarithmic error bound.

4. Main Empirical Result on Market-1501-τ

The dominant leak is exposure expansion. From R1 to R2, retrieval attacker A0 jumps from $P_{\text{guess}} = 0.288$ to 0.749 and $vRPI_2 = 4.332$ to 6.212. A3 rises from 0.407 to 0.706. R4 then shows that publishable artifacts themselves create exploitable leakage pathways.

Release	Attacker	P_{guess}	$vRPI_2$	Interpretation
R1	A0	.2883	4.3319	Sparse exposure baseline
R2	A0	**.7491**	**6.2119**	Dominant exposure-expansion jump
R3	A0	.7529	5.9271	Generic channel; calibrated fusion
R4	A2	**.7519**	**6.3572**	Prototype/index specialist
R4	A0-Graph	.7366	5.6697	kNN-graph pathway

Mean over 3 seeds. A3 independently shows the same R1→R2 pattern (P_{guess}: .407→.706; $vRPI_2$: 4.723→5.940). Empirical plug-in estimates can fluctuate with attacker/calibration even though Theorem 1 is a true-posterior statement.

5. Release Trend: The R1→R2 Jump Dominates

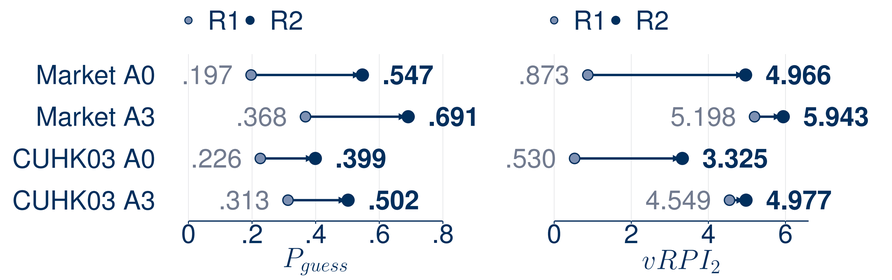
	R1	R2	R3	R4
A0	.2883	.7491	.7529	–
A1	.1600	.5375	.5372	–
A3	.4073	.7063	.7092	–
A4	.5073	.7304	.7238	–
A2	–	–	–	.7519
A0-Graph	–	–	–	.7366

	R1	R2	R3	R4
A0	4.3319	6.2119	5.9271	–
A1	1.6561	4.9997	4.9917	–
A3	4.7231	5.9402	5.9368	–
A4	5.0912	6.0133	6.0060	–
A2	–	–	–	6.3572
A0-Graph	–	–	–	5.6697

low high

Darker = higher (ramp scaled independently per metric). Blank cells mean no value is reported in the paper for that attacker/stage, not zero. The exposure expansion is the clearest compositional effect; R2→R3 stays in the same dark band under validation-tuned fusion and temperature scaling; R4 demonstrates artifact-specific leakage through the prototype/index specialist and graph-exploiting attacker.

6. Replication + Non-Face Stress Test



Light dot = R1, dark dot = R2; every arrow points right. Row order matches the paper's replication table.

CUHK03-τ replication. A lighter paired protocol still yields a positive R1→R2 increase for both retrieval (A0) and nonlinear-probe (A3) attackers on Market and CUHK03. The absolute magnitudes are not compared across protocols; the paired direction is the claim.

Masking ablation. Sweeping top-region masking from 0% to 40% leaves pressure essentially flat: A0 $vRPI_2 \in [4.87, 4.91]$, A3 $vRPI_2 \in [5.83, 5.88]$. Even at 40% masking, $P_{\text{guess}} \ge .51$ (A0) and $\ge .64$ (A3).

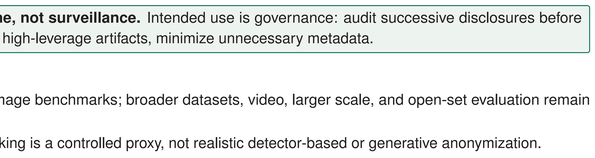
Important scope. The top-region mask is a controlled proxy, not a realistic detector-based or generative anonymizer. The result shows persistent non-face signal under aggressive region removal; it does not claim that all anonymization methods fail.

7. Experimental Protocol and Attacker Ladder

Dataset. Main experiment: Market-1501 training identities, $N = 751$, with the top 20% of each image masked (Market-1501-τ). Images are split into exposure, gallery, query, and held-out validation pools over the same closed-world label space.

What counts as a release?

- **R1:** encoder E_1 trained on sparse exposure.
- **R2:** add E_2 trained on full exposure; cumulative feature state.
- **R3:** add generic ImageNet-pretrained ResNet-50 E_3.
- **R4:** publish an identity prototype bank and a kNN graph derived from the R3 gallery.



Attacker ladder.

Every attacker maps scores into a posterior over the same identity universe before entropy computation.

8. Diagnostics: Dependence, Calibration, Priors

Dependence overlap. Conditional independence gives an ideal additive benchmark, not a claim about real CV releases. On Market,

$$\Delta_{\text{corr},2} = \sum_t \widehat{vRPI}_2(Z_t) - \widehat{vRPI}_2(Z_{1:T}) = 0.304 > 0,$$

quantifying redundancy among correlated release channels.

Temperature calibration. Across $\theta \in \{.025, .05, .1, .2, .5, 1.0\}$, absolute pressure shifts but $\Delta vRPI_2(R2-R1)$ stays positive throughout: A0 spans $[.781, 4.100]$, A3 spans $[.567, .677]$.

Non-uniform identity priors. For a Zipf(s) prior, normalizing by prior entropy instead of $\log N$, the R1→R2 increment shrinks from 3.818 at $s = 0$ to $.123$ at $s = 1.5$: a concentrated prior leaves less anonymity to erode.

9. What the Framework Adds

Not another Re-ID model. $vRPI_\alpha$ audits release-induced privacy pressure; rank/mAP remains a retrieval-utility metric – different questions.

Not a replacement for maximal leakage or DP. Maximal leakage optimizes adversarial gain per channel; DP gives worst-case stability guarantees. $vRPI_\alpha$ fixes the identity task and measures posterior concentration over a release sequence.

Governance implication. Privacy review should happen at the **sequence level**: evaluate planned disclosures together, minimize unnecessary metadata, and treat embedding dumps, prototype banks, and indices as privacy-relevant releases.

Positioning against common criteria.

Measure	Object measured	Sequence-aware?
Rank/mAP	Retrieval ranking quality	No
P_{guess}	Bayes posterior maximum	Yes, on $Z_{1:T}$
$vRPI_\alpha$	Rényi posterior concentration	**Yes**
Maximal leakage	Worst-case channel gain	Channel-level
DP / attribute inference	Worst-case stability / bound	Via composition

Release hygiene, not surveillance. Intended use is governance: audit successive disclosures before publication, limit high-leverage artifacts, minimize unnecessary metadata.

Limitations.

- Closed-world image benchmarks; broader datasets, video, larger scale, and open-set evaluation remain future work.
- Top-region masking is a controlled proxy, not realistic detector-based or generative anonymization.
- $vRPI_\alpha$ is an audit statistic, not a mitigation method.

10. Final Contribution, Empirical Signal, and Limitation

Contribution

A **release-sequence privacy formalism**, $vRPI_\alpha$, from Arimoto conditional Rényi entropy.

Theory: monotonicity under added side information, a Bayes-guessing bridge, an additive benchmark, ablation monotonicity, and plug-in stability.

Empirical signal

On face-masked Market-1501, **exposure expansion dominates**: A0 P_{guess} .288→.749, $vRPI_2$ 4.332→6.212.
CUHK03 preserves the paired R1→R2 direction; prototype/index and kNN-graph releases add artifact-specific leakage.

Limitation

Evaluation is theory-aligned and closed-world. $vRPI_\alpha$ is an **audit statistic, not a mitigation**.
Next: **open-set posteriors, realistic anonymizers, more datasets/video, and longer real release sequences.**