Does the Source of Carrier Image Affect Steganographic Detectability?

Midway Proposal – 5-Minute Version

Abdul Moiz Akbar | Malo Coquin | Daria Gjonbalaj | Nico Muller-Spath
Jimena Narvaez del Cid | David Wicker | Nikolas Zouros

Department of Advanced Computing Sciences
Maastricht University

Project 2.2 | February 2026

Agenda

1. Motivation and problem statement
2. State of the art and research gap
3. Research questions
4. Study design and methods
5. Experiment and evaluation plan
6. Prototype status
7. Research positioning and expected contribution
8. Minimum deliverable threshold

Motivation and Problem Statement

- Steganalysis detectability depends on carrier-image statistics, not only on embedding logic (Petitcolas et al., 1999; Cheddad et al., 2010; Fridrich & Kodovsky, 2012).
- Most benchmarks assume real camera images; this assumption is weaker as synthetic images become common.
- Diffusion and GAN images exhibit different statistical traces (Wang et al., 2020; Corvi et al., 2023).

Core problem

Do detectors validated on photographs remain equally effective on ML-generated carriers under identical embedding settings?

Study scope

Controlled $2 \times 2 \times 3 \times 2$ design, 1,000 images (500 real + 500 ML-generated), two embedding methods, three payload levels, two primary detectors.

State of the Art and Research Gap

- Classical post-hoc embedding and detection are well established (LSB, DCT-QIM, RS, χ^2, SRM).
- Generative and coverless steganography usually embed during generation, which is a different threat model than ours.
- Closest prior (De et al., 2022) shows feasibility on AI-generated images but not a controlled real-vs-ML detectability comparison under matched LSB/DCT settings.
- Synthetic-image forensics shows that AI-generated images carry distinct statistical traces, which motivates our carrier-origin hypothesis.

Research gap

No prior study in our scope isolates *carrier origin* as the key variable while keeping embedding pipeline, payload, and detector settings fixed.

Research Questions

RQ1 (Carrier Origin)

Is there any effect of carrier-image origin (real vs. ML-generated) on detectability of hidden data?

RQ2 (Payload)

Does increasing payload size widen the detectability gap between real and ML-generated carriers?

RQ3 (Encryption)

Does encrypting payload before embedding make steganography harder or easier to detect, and does origin change this effect?

RQ4 (Embedding Method)

Do different embedding methods (spatial LSB vs. frequency-domain DCT) interact differently with carrier origin in terms of detectability?

RQ5 (Image Quality)

Is image quality affected by embedding method and payload size?

Chosen Approaches: Data and Design

Factorial design

- Carrier type: real vs ML-generated
- Embedding: LSB vs DCT-QIM
- Payload: low, medium, high
- Detectors: RS, SRM+FLD

Datasets

- Real: RAISE (250), COCO (150), Flickr30k (100)
- ML-generated: Stable Diffusion v2.1 (250), StyleGAN3 (250)
- All images normalized to 512x512 RGB PNG
- BRISQUE $\leq$ 50 quality gate for generated images

Chosen Approaches: Methods and Validation

Embedding methods

- LSB substitution in spatial domain ($k = 1, 2$)
- DCT-QIM in frequency domain (8x8 blocks; zigzag 10–54)
- Payload optionally encrypted with AES-256-CBC

Detection and validation

- Primary detectors: RS Analysis, SRM+FLD
- Supplementary check: χ^2 attack (LSB)
- Metrics: ROC-AUC, EER, accuracy at Youden's J
- Quality: PSNR, SSIM, FSIM

Experiment and Evaluation Plan

Experiment	Focus
Exp.1 (RQ1)	Compare real vs. ML detectability under matched settings.
Exp.2 (RQ2)	Test whether payload level widens the real-vs-ML AUC gap.
Exp.3 (RQ3)	Compare plain vs. AES-encrypted payload detectability.
Exp.4 (RQ4)	Run two-way ANOVA for carrier origin × embedding method.
Exp.5 (RQ5)	Assess image quality (PSNR/SSIM/FSIM) across conditions.

Statistical reporting

We evaluate performance primarily with ROC-AUC and report uncertainty and effect size alongside significance, using a stricter threshold to account for multiple comparisons.

Prototype Status

Vertical prototype

- LSB embedding/extraction
- DCT-QIM embedding
- RS Analysis
- SRM feature extraction/classification

Horizontal prototype

- Integrated run on 50 images
- 25 real + 25 ML-generated
- Medium-payload pipeline check
- Interfaces validated before full run

Research Positioning and Expected Contribution

What we are trying to do

- Quantify whether ML-generated carriers change detectability.
- Compare real vs ML under identical LSB/DCT embedding conditions.
- Measure effects of payload and encryption, not just feasibility.

How it fits current research

- Bridges steganalysis and synthetic-image forensics.
- Uses standardized classical baselines for comparability.
- Produces direct evidence on whether existing detectors transfer to mixed real/synthetic traffic.

Minimum Deliverable Threshold

Approach threshold

- At least one encryption algorithm
- At least two embedding methods:
- one spatial-domain and one frequency-domain

Research threshold

At minimum, answer:

- RQ3 (Encryption)
- RQ4 (Embedding-method interaction)

Thank you

Questions and Discussion

Document: `docs/proposals/midway_proposal_final.tex`