Computational Discovery and Intelligent Systems 5(2026):1-55

Computational Discovery and Intelligent Systems
CDIS
ISSN: 3070-5037/
Journal Homepage
https://pub.scientificirg.com/index.php/CDIS

A Robust Cross-Sensor Image Steganalysis Framework for Robotic Telemetry in Nuclear Facilities

Mostafa Morsy Mohamed [a], Mohamed Ragab Salama [a], and Ayat Taha [a, b,1]

[a] *ICT Department, New Cairo Technological University (NCTU), Cairo, Egypt;*
E-mails: 243101@nctu.edu.eg, mohamedsalama1162000@gmail.com, and ayat.taha.st@nctu.edu.eg
[b] *BT Department, Egyptian Russian University, Cairo, Egypt; E-mail: ayat-taha@eru.edu.eg*

ABSTRACT

Autonomous robotic platforms increasingly support inspection, surveillance, and environmental monitoring at safety-critical facilities. Image telemetry from these platforms creates a potential covert channel because a payload can be embedded in otherwise legitimate frames before transmission. This study evaluates a convolutional steganalysis network, NuclearStegNet, that combines fixed Spatial Rich Model residual filters with residual attention blocks and Squeeze-and-Excitation channel weighting. The evaluation uses BOSSbase 1.01 as an optical reference source and HIT-UAV thermal imagery as a cross-sensor test source. On the supplied experimental record, NuclearStegNet reaches 89.3% accuracy for J-UNIWARD at 0.4 bits per pixel on the matched optical test condition and 82.8% under optical-to-thermal mismatch. The corresponding SRM with Ensemble Classifier baseline reaches 68.4% and 59.9%, respectively. These findings suggest that residual preprocessing and channel attention may enhance detection under the evaluated distribution shift. However, the thermal dataset is not representative of nuclear inspection imagery and cannot establish operational performance in a nuclear facility. Accordingly, the study is presented as a reproducibility focused research framework with a preliminary deployment concept, rather than as evidence of an observed attack or a validated field system.

PAPER INFORMATION

HISTORY
Received: 16 May 2026
Revised: 19 June 2026
Accepted: 18 July 2026
Online: 29 August 2026

MSC
68T07; 68R10; 94A60; 68M15

KEYWORDS
Image Steganalysis;
Robotic Telemetry;
Channel Attention;
HIT-UAV;
J UNIWARD.

1 INTRODUCTION

Digital imagery is an important information product in robotic inspection, perimeter surveillance, and environmental monitoring. In a safety-critical setting, the image stream may contain operationally sensitive information even when the associated network connection is protected by standard authentication and encryption. Steganography creates a different security concern from direct network intrusion because a sender can alter the statistical content of an apparently valid image while preserving its ordinary file-level appearance. Detection of the embedding operation is therefore a complementary problem to encryption, access control, and network monitoring. Modern adaptive schemes allocate modifications according to local image complexity. UNIWARD-based and directional-filtering methods reduce distortion in textured or noisy regions, which can make detection more difficult than detection of simple least-significant-bit replacement [1, 2, 3]. The relevant concern in robotic telemetry is not that a particular facility has experienced such an attack. Rather, an image stream can constitute a plausible covert channel when an adversary has access to a camera pipeline or an intermediate processing

[1]Corresponding author at BT Department, Egyptian Russian University, Cairo, Egypt; E-mail: ayat-taha@eru.edu.eg

component. The present study treats this possibility as a hypothetical threat model and does not claim an observed incident, a recovered payload, or a compromise of a nuclear control system. The assumed adversary has temporary logical access to an image-processing component on an autonomous robotic platform and can apply a content-adaptive embedding algorithm to outbound image frames. The communication channel, underlying control system, authenticated encryption, and hardware security modules remain uncompromised. The detector is intended to identify whether embedding has occurred; payload recovery, operator identification, device attribution, and confirmation of a complete attack chain are outside its scope. Most steganalysis benchmarks are based on photographic images. BOSSbase and related datasets have supported systematic comparisons of rich models, ensemble classifiers, and deep residual architectures [4, 5, 6, 7]. Robotic inspection platforms, by contrast, may rely on thermal sensors, nonstandard optical systems, radiometric transformations, compression procedures, and scene conditions that differ substantially from those represented in photographic benchmarks. Consequently, a detector trained on one image source may learn source-specific characteristics rather than embedding artifacts. This phenomenon, commonly referred to as cover-source mismatch, represents a significant challenge for the reliable deployment of steganalysis systems [8, 9]. This article evaluates NuclearStegNet under matched and cross-sensor conditions. The architecture uses 30 fixed residual filters followed by five residual attention blocks with Squeeze-and-Excitation channel weighting [10]. The reported experiments compare the proposed model with the SRM and Ensemble Classifier baseline, SRNet, and Yedroudj Net. The comparison is situated within the progression from learned hierarchical residual representations to deeper residual architectures [11, 7]. The analysis also documents the partitioning rule, embedding-quality measurements, cross-sensor protocol, and limits of interpreting HIT-UAV as a proxy for nuclear inspection imagery. The principal contribution is a clearly specified framework for studying image steganalysis under sensor mismatch, together with an explicit separation between reported benchmark performance and unvalidated field deployment. The contribution is not a claim that the model is ready for safety-critical deployment. The remainder of this paper is organized as follows. Section 2 reviews related work in adaptive steganography and steganalysis. Section 3 describes the NuclearStegNet architecture and training protocol. Section 4 details the experimental datasets and partitioning rules. Section 5 presents the matched and cross-sensor results. Section 6 discusses operational calibration and deployment considerations. Section 7 addresses limitations, and Section 8 concludes the study.

2 RELATED WORK

This section first reviews adaptive embedding methods, then summarizes representative steganalysis architectures, and finally positions the experiment within the cover-source mismatch literature.

2.1 Adaptive Image Steganography

Classical image steganography includes spatial replacement, pixel value differencing, and transform domain modification. Adaptive methods instead minimize a distortion objective that depends on local image structure. UNIWARD methods define distortion using wavelet domain changes, while WOW uses directional residual responses to assign embedding costs [1, 2, 3]. These methods are useful test cases because their modifications are designed to avoid simple statistical tests.

2.2 Steganalysis Models

Rich models aggregate residual statistics generated by a large bank of high-pass filters and have been combined with ensemble classifiers for reliable detection on matched sources [5, 6]. Early convolutional approaches demonstrated that hierarchical representations can be learned directly for image steganalysis [11]. Later residual networks improved the depth and optimization of the learned detector while retaining a constrained residual front end [7]. Yedroudj Net introduced an efficient convolutional design for spatial steganalysis [12]. The ALASKA challenge demonstrated the importance of broad source diversity and realistic processing variation in evaluating generalization [13]. Selection channel information and repeated embedding keys also affect detector performance [14, 15].

The studies summarized above establish the methodological context for comparing adaptive embedding algorithms, residual models, and cross-source evaluation. The present work applies that context to a controlled optical-to-thermal proxy setting without claiming generalization to all sensors or operational facilities.

Table 1 summarizes the representative methods reviewed in this section.

Table 1: Summary of representative steganography and steganalysis studies

Study or model	Main contribution	Limitation relevant to this study
UNIWARD and WOW [1, 2, 3]	Content-adaptive distortion and directional embedding costs	Detection difficulty depends on source, payload, and processing conditions
Rich models and ensemble classifiers [5, 6]	Hand-designed residual statistics with ensemble decision rules	Performance can degrade under source mismatch
YeNet [11]	Learned hierarchical representations for image steganalysis	Requires evaluation beyond the training distribution
SRNet [7]	Deep residual architecture with a constrained residual front end	Generalization remains source-dependent
Yedroudj Net [12]	Efficient convolutional architecture for spatial steganalysis	Primarily established on photographic sources
CSM studies [8, 9]	Analysis of cover-source mismatch and possible mitigation strategies	Do not validate the present optical-to-thermal experiment

2.3 Cover-Source Mismatch

Cover-source mismatch arises when training and testing images differ in camera, scene content, resizing, compression, or other processing factors. A detector can then respond to the source difference rather than to the embedding signal. Recent studies describe CSM as a central operational problem and examine domain adaptation, source diversity, and calibration as possible remedies [8, 9]. The present evaluation focuses on a deliberately difficult optical-to-thermal shift and is interpreted as a controlled proxy experiment.

3 PROPOSED METHOD

3.1 NuclearStegNet Architecture

NuclearStegNet combines fixed residual extraction with learned channel attention. Let $x \in \mathbb{R}^{H\times W}$ denote a single channel image and let $h_k \in \mathbb{R}^{5\times 5}$ denote the k-th fixed SRM filter. The residual response for filter k is defined by **Equation 1**.

$$r_k(i,j) = \sum_{u=-2}^{2}\sum_{v=-2}^{2} h_k(u,v)\,x(i-u,j-v), \qquad k \in \{1,\ldots,30\}. \tag{1}$$

The residual representation is consistent with rich-model steganalysis [5]. The boundary handling rule must be fixed in the implementation, for example by symmetric padding, because different padding rules change the residual distribution. The residual stack is $R = [r_1,\ldots,r_{30}] \in \mathbb{R}^{30\times H\times W}$. Fixed filters constrain the first representation to emphasize local prediction error rather than scene semantics.

Five residual attention blocks follow the filter layer. Each block contains two 3×3 convolutions, batch normalization, rectified linear activation, a skip connection, and a Squeeze-and-Excitation module. For an input feature tensor $U \in \mathbb{R}^{C\times H\times W}$, global average pooling produces the channel descriptor $z \in \mathbb{R}^{C}$, whose c-th component is defined by **Equation 2**.

$$z_c = \frac{1}{HW}\sum_{i=1}^{H}\sum_{j=1}^{W} U_{cij}, \qquad c \in \{1,\ldots,C\}. \tag{2}$$

With reduction ratio $r = 16$, the channel gate is defined by **Equation 3**.

$$s = \sigma(W_2\,\delta(W_1 z)), \quad W_1 \in \mathbb{R}^{C/r\times C}, \quad W_2 \in \mathbb{R}^{C\times C/r}, \tag{3}$$

where $\delta(a) = \max(0,a)$ is applied elementwise and $\sigma(a) = 1/(1+e^{-a})$ is the elementwise sigmoid function [10]. The reweighted feature tensor is defined by **Equation 4**.

$$\widetilde{U}_{cij} = s_c U_{cij}. \tag{4}$$

Global average pooling is then applied to the final feature tensor, followed by fully connected layers with dimensions 256, 128, and 2. Dropout with probability 0.5 is applied before the final classifier. NuclearStegNet utilizes approximately 5.2 million trainable parameters according to the reported experimental configuration.

Figure 1 presents the proposed processing sequence before the detailed training protocol.

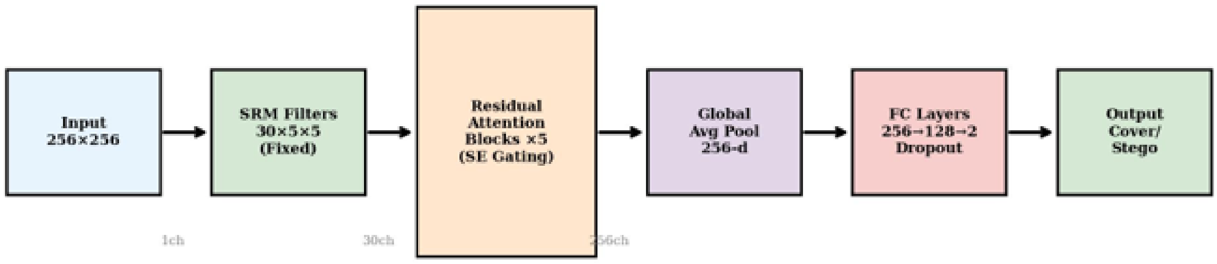

Figure 1: NuclearStegNet architecture with fixed residual filtering and channel attention

3.2 *Training Protocol*

Training uses Adam with an initial learning rate of 10^{-3}, weight decay of 10^{-4}, cosine annealing over 200 epochs, and a batch size of 32. For a mini-batch of N samples with binary labels $y_n \in \{0, 1\}$ and predicted stego probabilities p_n, the cross-entropy objective is given by **Equation 5**.

$$\mathcal{L}_{\text{CE}} = -\frac{1}{N} \sum_{n=1}^{N} \left[y_n \log p_n + (1 - y_n) \log(1 - p_n) \right] . \tag{5}$$

Random horizontal and vertical flips, rotations by multiples of 90 degrees, and random crops are applied only to the training partition. This objective is minimized during training. The SRM layer is frozen during the reported thermal adaptation stage, while later layers are fine-tuned for 50 epochs.

3.3 *Evaluation and Reproducibility*

A rigorous evaluation should treat the cover image, rather than the individual crop or stego derivative, as the unit of partitioning. This rule prevents correlated cover and stego pairs from appearing in different partitions and avoids optimistic estimates caused by near duplicate patches. Training and validation transformations should be fitted only to the training data, and test images should remain untouched until the final evaluation.

Accuracy alone is insufficient for an operational detector because the cost of false alarms can be high. Let TP, TN, FP, and FN denote the numbers of true positives, true negatives, false positives, and false negatives at decision threshold τ. The true-positive and false-positive rates are defined by **Equation 6**.

$$\text{TPR}(\tau) = \frac{TP(\tau)}{TP(\tau) + FN(\tau)}, \qquad \text{FPR}(\tau) = \frac{FP(\tau)}{FP(\tau) + TN(\tau)}, \tag{6}$$

Precision and balanced accuracy are defined by **Equation 7**.

$$\text{Precision}(\tau) = \frac{TP(\tau)}{TP(\tau) + FP(\tau)}, \qquad \text{BA}(\tau) = \frac{\text{TPR}(\tau) + 1 - \text{FPR}(\tau)}{2}. \tag{7}$$

The predicted class is defined by $\widehat{y} = 1$ when $p \geq \tau$ and by $\widehat{y} = 0$ otherwise. The evaluation should report these measures together with the areas under the receiver operating characteristic and precision recall curves. Threshold selection should be performed on the validation partition and then frozen before test evaluation. Calibration should be assessed with reliability diagrams and a proper scoring rule such as the Brier score. A threshold selected after inspecting the test predictions would produce an optimistic estimate and should not be used.

4 DATA AND EXPERIMENTAL PROTOCOL

4.1 *Dataset Composition and Selection*

The evaluation uses two primary image sources, summarized in **Table 2**. BOSSbase 1.01 provides 10,000 grayscale photographic images captured by seven consumer digital single lens reflex cameras [4]. HIT-UAV is a high altitude infrared thermal dataset containing 2,898 frames captured by a DJI Matrice 300 RTK platform with a Zenmuse H20T thermal camera operating in the long wave infrared band from 8 to 14 micrometers [16]. HIT-UAV serves as a public thermal proxy in this study and is not a nuclear inspection dataset. From each primary source, 1,000 cover images are selected for the reported experiments; the exact image identifiers and the sampling procedure are recorded in the accompanying reproducibility package. A third ground based thermal source using FLIR and Seek sensors is used only for descriptive sensor statistics and is not included in the detection experiments.

Table 2: Dataset roles and quantities in the reported evaluation

Source	Original source description	Experimental use	Quantity used
BOSSbase 1.01	Grayscale photographic images from seven DSLR cameras	Optical training, validation, and testing	1,000 selected covers
HIT-UAV	Infrared thermal images extracted from UAV video frames	Cross-sensor and thermally matched evaluation	1,000 selected covers
Ground thermal source	FLIR and Seek imagery described in the supplied record	Descriptive statistics only	Not used for detection

Table 3 reports the descriptive statistics used to contextualize the source shift.

Table 3: Descriptive statistics for the evaluated sensor sources

Metric	BOSSbase	HIT-UAV	Ground FLIR
Mean intensity	124.2	141.9	138.2
Standard deviation	47.8	31.6	26.8
Dynamic range	255	254	224
Entropy (bits)	7.58	7.02	6.78
Skewness	0.01	0.01	-0.01

The numerical values correspond to **Figure 2**. Because the source record does not specify the complete sampling and calculation procedure, the table is not used to support a significance claim.

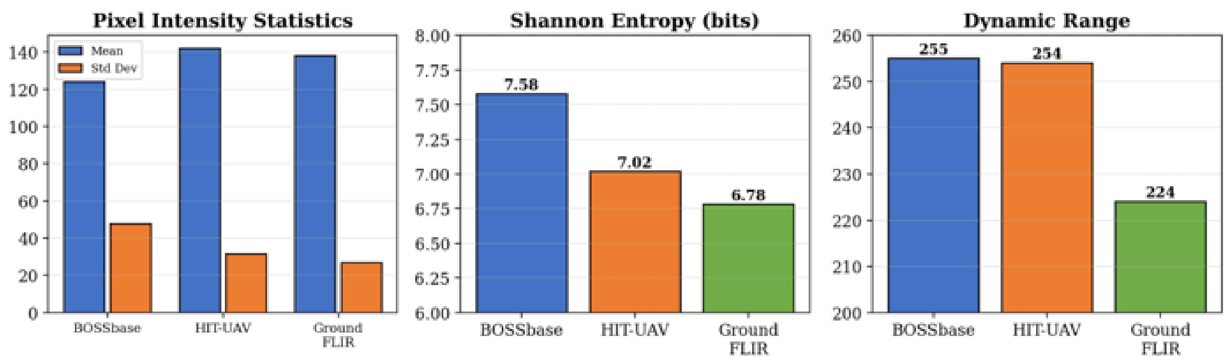

Figure 2: Descriptive sensor statistics from the supplied record

4.2 Embedding and Partitioning Protocol

Each selected cover is paired with a stego image generated at payload rates of 0.1, 0.2, and 0.4 bits per pixel (bpp) using J-UNIWARD, S-UNIWARD, and WOW. Cover images are assigned to training, validation, or test partitions before stego generation. All stego derivatives of a cover remain in the same partition to prevent data leakage. The supplied split is 60% for training (600 covers), 20% for validation (200 covers), and 20% for testing (200 covers) with random seed 42. No crop or derivative of a cover is permitted to cross a partition boundary.

Three evaluation conditions are considered. The matched optical condition trains and tests on BOSSbase, the cross-sensor condition trains on BOSSbase and tests on HIT-UAV, and the matched thermal condition trains and tests on HIT-UAV. Further independent evaluation would benefit from reporting the number of unique scenes, the exact image identifiers, preprocessing code, embedding implementation, random seeds, and per-run confidence intervals.

4.3 *Quality Measures*

Peak signal-to-noise ratio and structural similarity are computed between each cover and its stego counterpart. Let x and $\widehat{x}$ denote the cover and stego images, respectively, with $M = HW$ pixels. The mean-squared error is defined by **Equation 8**.

$$\mathrm{MSE}(x, \widehat{x}) = \frac{1}{M} \sum_{i=1}^{H} \sum_{j=1}^{W} \left[x_{ij} - \widehat{x}_{ij}\right]^2 . \tag{8}$$

For a maximum representable intensity L, peak signal-to-noise ratio is defined by **Equation 9**.

$$\mathrm{PSNR}(x, \widehat{x}) = 10 \log_{10}\left(\frac{L^2}{\mathrm{MSE}(x, \widehat{x})}\right). \tag{9}$$

The value of L must be stated with the image bit depth and preprocessing convention. Structural similarity (SSIM) should likewise be computed with a declared window size, dynamic range, and implementation, because these choices affect the reported value. For local means μ_x and $\mu_{\widehat{x}}$, local variances σ_x^2 and $\sigma_{\widehat{x}}^2$, local covariance $\sigma_{x\widehat{x}}$, and stabilizing constants C_1 and C_2, the standard two-component SSIM expression is given by **Equation 10**.

$$\mathrm{SSIM}(x, \widehat{x}) = \frac{(2\mu_x\mu_{\widehat{x}} + C_1)(2\sigma_{x\widehat{x}} + C_2)}{(\mu_x^2 + \mu_{\widehat{x}}^2 + C_1)(\sigma_x^2 + \sigma_{\widehat{x}}^2 + C_2)}. \tag{10}$$

The SSIM framework compares luminance, contrast, and structural information between the reference and distorted images [17]. The expression should be evaluated over the declared local windows and then averaged according to the selected implementation. **Figure 3** provides the corresponding visual comparison.

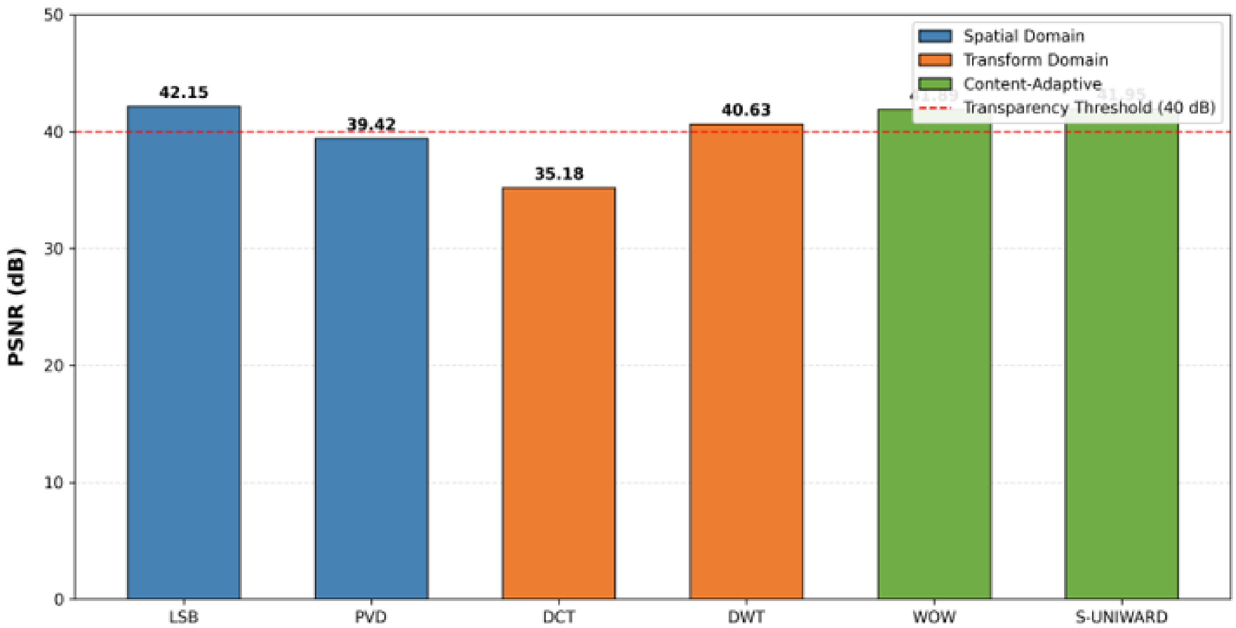

Figure 3: PSNR comparison at 0.4 bpp

Table 4 summarizes the reported PSNR and SSIM values at the 0.4 bpp payload condition, while Figure 3 provides the corresponding PSNR comparison.

Table 4: Embedding quality at 0.4 bpp

Method	Domain	PSNR (dB)	SSIM	bpp
LSB	Spatial	42.15	0.9854	0.4000
PVD	Spatial	39.42	0.9712	0.4000
DCT	Transform	35.18	0.9425	0.4000
DWT	Transform	40.63	0.9791	0.4000
WOW	Adaptive	41.89	0.9912	0.4000
S-UNIWARD	Adaptive	41.95	0.9924	0.4000

5 RESULTS

This section presents the results in three stages. Matched optical performance is reported first, followed by the cross-sensor comparison and then the ablation and receiver operating characteristic analysis. The reported values are interpreted as descriptive because the supplied record does not include the complete repeated-run evidence required for confirmatory inference.

5.1 Matched Optical Detection

The matched experiment compares the proposed model with the SRM with Ensemble Classifier baseline across three payload rates and three embedding algorithms. The proposed model improves over the baseline in every listed condition. At 0.4 bits per pixel for J-UNIWARD, the reported accuracy is 89.26% for NuclearStegNet and 68.42% for the baseline, an absolute difference of 20.84 percentage points. The corresponding 89.3% value in the figure is the rounded presentation value.

Table 5 gives the matched optical results across the three payload rates.

Table 5: Matched optical detection accuracy

Algorithm	Payload	0.1		0.2		0.4	
		NSN	SRM	NSN	SRM	NSN	SRM
J-UNIWARD		72.14	54.82	79.56	60.15	89.26	68.42
S-UNIWARD		73.55	57.11	81.84	62.90	90.48	70.85
WOW		74.92	58.64	83.10	64.38	92.23	72.91

Figure 4 visualizes the principal matched optical comparison.

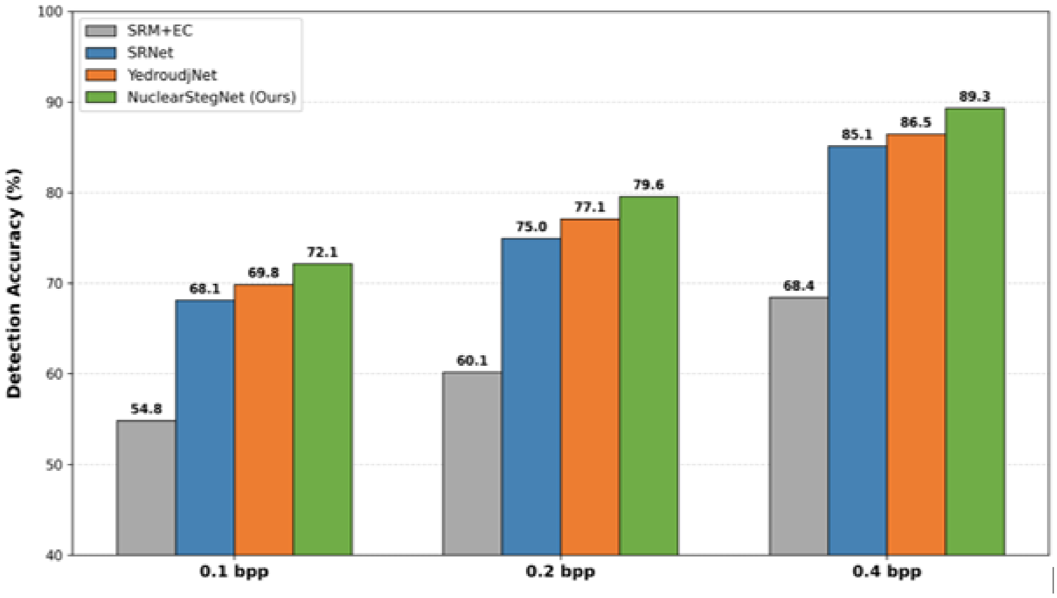

Figure 4: Matched optical accuracy for J-UNIWARD

5.2 Cross-Sensor Detection

Under the optical-to-thermal condition, NuclearStegNet reaches 82.79% and the SRM with Ensemble Classifier reaches 59.85% at 0.4 bits per pixel for J-UNIWARD. The absolute reduction from the matched optical condition is 6.47 percentage point for NuclearStegNet and 8.57 percentage points for the baseline. The thermal domain matched result for NuclearStegNet is 86.07%. These figures support the narrower conclusion that the proposed architecture was less affected by the tested source shift than the listed baseline. These values do not establish general robustness across thermal cameras, radiometric preprocessing pipelines, or nuclear facility scenes.

Table 6 reports the numerical cross-sensor comparison before the corresponding figure.

Table 6: Cross-sensor results for J-UNIWARD at 0.4 bpp

Condition	Sources	SRM+EC	SRNet	Yedroudj Net	NSN
Matched optical	BOSSbase/BOSSbase	68.42	85.12	86.45	89.26
Optical-to-thermal	BOSSbase/HIT-UAV	59.85	77.92	78.60	82.79
Matched thermal	HIT-UAV/HIT-UAV	64.10	81.25	82.51	86.07

Table 7 consolidates the reported operating-point results for direct comparison.

Table 7: Consolidated comparison at the reported operating points

Method	Parameters	BOSS 0.2	BOSS 0.4	Thermal CSM	CSM drop
SRM+EC	34.7K	60.15	68.42	59.85	8.57
SRNet	3.8M	74.95	85.12	77.92	7.20
Yedroudj Net	4.1M	77.10	86.45	78.60	7.85
NuclearStegNet	5.2M	79.56	89.26	82.79	6.47

This consolidated table presents values that are otherwise distributed across the matched and cross-sensor tables. The table is not an additional experiment; it is a compact presentation of the supplied operating-point results. The table consolidates the reported operating-point results for direct comparison across the evaluated methods.

Table 8 records the exploratory aggregate summary and its limited inferential status.

Table 8: Exploratory cross-algorithm summary at 0.4 bpp

Metric	NuclearStegNet	SRM+EC
Mean accuracy across three algorithms	90.9 ± 1.7%	71.3 ± 2.3%
Mean AUC	0.962 ± 0.008	0.891

5.3 Ablation and ROC Analysis

The supplied ablation results attribute 3.7 percentage points to Squeeze-and-Excitation attention, 2.2 percentage points to fixed rather than learned residual filters, 2.5 percentage points to the fifth residual block, and 6.1 percentage points to skip connections. The values are useful for architectural interpretation but should be confirmed with repeated training runs because a single reported result per ablation does not quantify optimization variability.

For J-UNIWARD at 0.4 bits per pixel, the supplied record reports an area under the receiver operating characteristic curve of 0.954 for NuclearStegNet, 0.891 for SRM with Ensemble Classifier, 0.938 for SRNet, and 0.932 for Yedroudj Net. Under the optical-to-thermal condition, the reported NuclearStegNet area under the curve is 0.912. False positive operating points are important for facility monitoring, and the supplied record reports a true positive rate of 87.1% at 5% false positive rate on matched data and 79.4% under mismatch.

Figure 5 visualizes the cross-sensor comparison.

The ablation values are descriptive because the source record does not provide repeated runs or uncertainty intervals.

Figure 6 presents the component ablation values discussed above.

6 OPERATIONAL CALIBRATION AND DEPLOYMENT CONSIDERATIONS

A deployment threshold should be selected using a sensor-specific calibration set that is separate from the final test set. The calibration set should contain ordinary frames collected under representative operating conditions, including expected changes in illumination, thermal background, compression, resizing, and mission phase. A fixed threshold should not be transferred between sensors without measuring the resulting false positive rate. Monitoring should also record the model version, preprocessing version, sensor identifier, decision threshold, and time of inference so that an alert can be reconstructed during incident review.

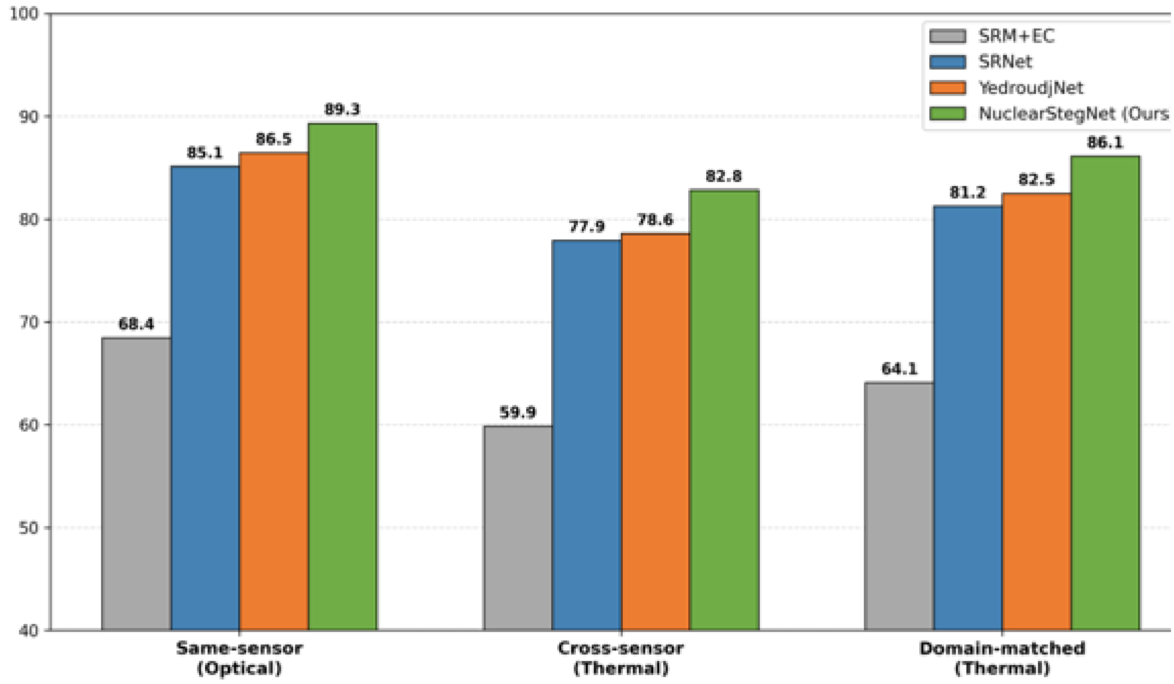

Figure 5: Cross-sensor detection results for J-UNIWARD at 0.4 bpp

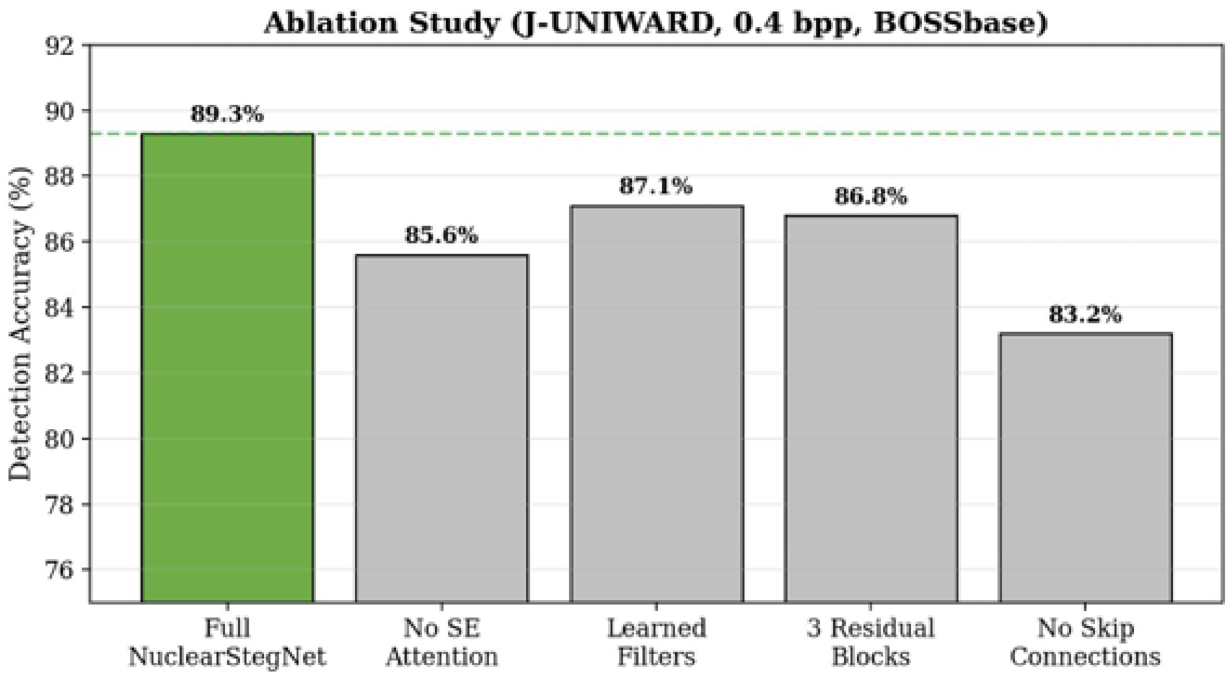

Figure 6: Ablation results for J-UNIWARD at 0.4 bpp

A practical monitoring system should use complementary controls rather than rely on a single classifier. Authenticated encryption protects confidentiality and integrity but does not by itself provide a content-based test for covert embedding by a trusted or compromised processing component. A gateway can therefore combine frame authentication, lightweight residual screening, server-side steganalysis, and behavioral monitoring. Each layer has a different failure mode and should generate evidence for incident review rather than an automatic assumption of compromise.

Authenticated encryption, sequence control, and key management remain primary controls. Residual screening can prioritize frames for more expensive analysis, while a trained detector can provide a second opinion at a protected gateway. Behavioral features such as unusual frame size, compression changes, transmission timing, or mission context can support triage but are not specific to steganography. A safe operational design should include calibration data from the deployed sensors, a human review path, versioned thresholds, and a documented procedure for false alarm analysis.

Figure 7 presents the conceptual defense-in-depth arrangement discussed in this section.

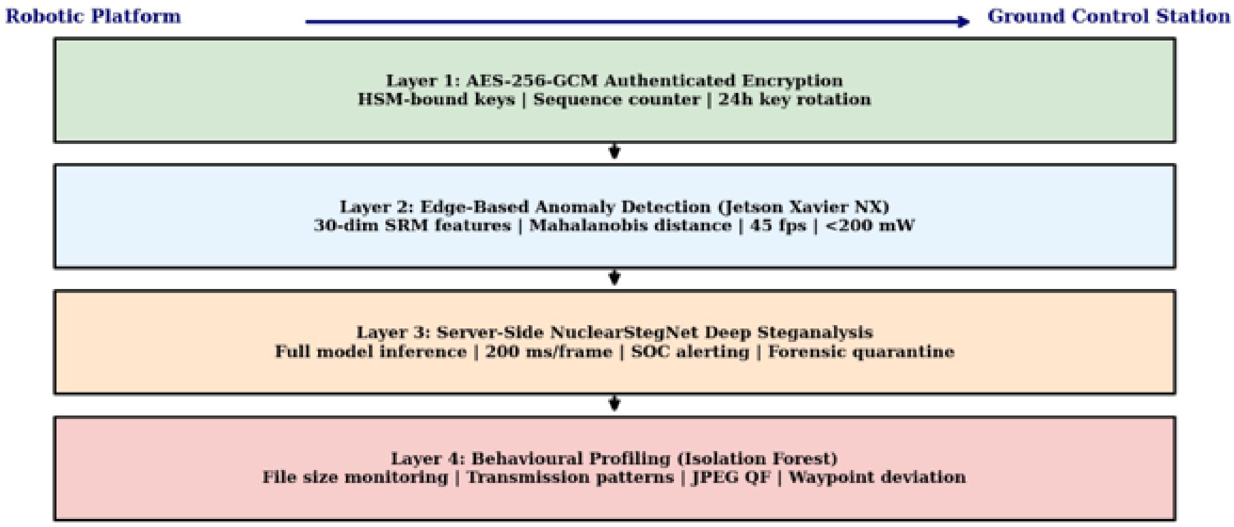

Figure 7: Conceptual four-layer defense-in-depth architecture

7 LIMITATIONS

The principal limitation is external validity. HIT-UAV is a public thermal dataset for object detection, not a nuclear inspection corpus. Differences in scene composition, sensor calibration, atmospheric conditions, radiometric encoding, compression, and operational geometry may change the detector response. The results therefore measure a controlled optical-to-thermal mismatch rather than performance in a reactor building, spent fuel pool, exclusion zone, or radiological environment.

The second limitation concerns statistical inference. Three embedding algorithms are not sufficient to support a broad claim of statistical significance across the steganalysis problem. The previously supplied sign-test value is not reported because the paired outcomes and analysis procedure are unavailable. Future evaluation should use repeated independent runs, image-level confidence intervals, bootstrap intervals for area under the curve, calibration assessment, and prespecified comparisons across payload rates and sources.

The third limitation is adversarial adaptation. A detector trained on known embedding families may fail against a new algorithm, a changed key, post-embedding recompression, or an adaptive adversary. Video telemetry is also outside the present scope. Temporal redundancy and interframe processing introduce additional embedding and detection mechanisms that cannot be represented by a still image classifier.

8 CONCLUSION

This study presents a cautious framework for detecting image steganography under a controlled cross-sensor shift. NuclearStegNet combines fixed residual filters with residual attention and reports higher accuracy than the listed classical and deep baselines on the supplied BOSSbase and HIT-UAV experiments. The reported reduction in performance under optical-to-thermal mismatch is consistent with a material effect of source shift, even when the detector performs well on the matched source. The results support research into sensor-specific calibration, source-diverse training, domain adaptation, and independent red-team evaluation of robotic image telemetry. The results do not demonstrate an observed nuclear-security incident, establish that HIT-UAV represents nuclear-inspection conditions, or justify deployment without a complete reproducibility package and field validation. A defense-in-depth monitoring design is therefore more appropriate than reliance on a single learned detector.

AUTHOR CONTRIBUTION STATEMENT

All authors contributed equally to the study conception and design. Material preparation, data collection, and analysis were performed by the authors. The first draft of the manuscript was written by the authors, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

This study did not involve human participants or animals. Ethical approval and consent to participate are therefore not applicable.

CONSENT FOR PUBLICATION

Not applicable.

DATA AVAILABILITY

This study uses the publicly available BOSSbase 1.01 and HIT-UAV image datasets. The analysis is based on selected subsets of these datasets and reports aggregate experimental results.

ACKNOWLEDGMENTS

The authors acknowledge the use of ChatGPT-5 for assistance in refining the manuscript. The authors reviewed and edited all AI-generated output and take full responsibility for the final content.

FUNDING

This research received no external funding.

DISCLOSURE STATEMENT

The authors declare that they have no competing interests.

REFERENCES

[1] V. Holub, J. Fridrich, and T. Denemark, "Universal distortion function for steganography in an arbitrary domain," *EURASIP Journal on Information Security*, vol. 2014, no. 1, p. 1, 2014.

[2] V. Holub and J. Fridrich, "Designing steganographic distortion using directional filters," in *2012 IEEE International workshop on information forensics and security (WIFS)*, pp. 234–239, IEEE, 2012.

[3] V. Holub and J. Fridrich, "Digital image steganography using universal distortion," in *Proceedings of the first ACM workshop on Information hiding and multimedia security*, pp. 59–68, 2013.

[4] P. Bas, T. Filler, and T. Pevnỳ, "Break our steganographic system": the ins and outs of organizing boss," in *International workshop on information hiding*, pp. 59–70, Springer, 2011.

[5] J. Fridrich and J. Kodovský, "Rich models for steganalysis of digital images," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 3, pp. 868–882, 2012.

[6] J. Kodovsky, J. Fridrich, and V. Holub, "Ensemble classifiers for steganalysis of digital media," *IEEE Transactions on information forensics and security*, vol. 7, no. 2, pp. 432–444, 2011.

[7] M. Boroumand, M. Chen, and J. Fridrich, "Deep residual network for steganalysis of digital images," *IEEE transactions on information forensics and security*, vol. 14, no. 5, pp. 1181–1193, 2018.

[8] G. Quentin, B. Patrick, C. Rémi, and B. Dirk, "The cover source mismatch problem in deep-learning steganalysis," in *2022 30th European Signal Processing Conference (EUSIPCO)*, pp. 1032–1036, IEEE, 2022.

[9] A. Mallet, M. Beneš, and R. Cogranne, "Cover-source mismatch in steganalysis: Systematic review," *EURASIP Journal on Information Security*, vol. 2024, no. 1, p. 26, 2024.

[10] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 7132–7141, 2018.

[11] J. Ye, J. Ni, and Y. Yi, "Deep learning hierarchical representations for image steganalysis," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 11, pp. 2545–2557, 2017.

[12] M. Yedroudj, F. Comby, and M. Chaumont, "Yedroudj-net: An efficient cnn for spatial steganalysis," in *2018 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pp. 2092–2096, IEEE, 2018.

[13] R. Cogranne, Q. Giboulot, and P. Bas, "The alaska steganalysis challenge: A first step towards steganalysis," in *Proceedings of the ACM Workshop on Information Hiding and Multimedia Security*, pp. 125–137, 2019.

[14] T. Denemark, V. Sedighi, V. Holub, R. Cogranne, and J. Fridrich, "Selection-channel-aware rich model for steganalysis of digital images," in *2014 IEEE international workshop on information forensics and security (WIFS)*, pp. 48–53, IEEE, 2014.

[15] L. Pibre, P. Jérôme, D. Ienco, and M. Chaumont, "Deep learning is a good steganalysis tool when embedding key is reused for different images, even if there is a cover source-mismatch," *arXiv preprint arXiv:1511.04855*, 2015.

[16] J. Suo, T. Wang, X. Zhang, H. Chen, W. Zhou, and W. Shi, "Hit-uav: A high-altitude infrared thermal dataset for unmanned aerial vehicle-based object detection," *Scientific Data*, vol. 10, no. 1, p. 227, 2023.

[17] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.